APR 21 1930

...DUCATIONAL ROOM
...ERAL LIBRARY
...NIV. OF MICH.

The Journal of EDUCATIONAL PSYCHOLOGY

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

HAROLD RUGG, *Editor*

VOL. XXI, No. 4 APRIL, 1930 $5.00 per year

CONTENTS

SEX DIFFERENCES: INTEREST IN TASKS REQUIRING MECHANICAL ABILITY AND MOTOR SKILL. *Harvey C. Lehman and Paul A. Witty*. 239

A STUDY OF MIND-SET IN RURAL AND CITY SCHOOL CHILDREN. *J. M. Shales*. 246

AN EXPERIMENTAL COMPARISON OF THE MOVEMENTS IN MANUSCRIPT WRITING AND CURSIVE WRITING. *William Henry Gray* 259

SHOULD TEACHERS GIVE WARNING OF TESTS AND EXAMINATIONS? *Glenn R. Pease*. 273

THE RESTANDARDIZATION OF IQ'S OF DIFFERENT TESTS. *Minnie L. Steckel* 278

THE INTERPRETATION OF CORRELATION COEFFICIENTS IN TERMS OF DEPARTURE FROM PERFECT CORRELATION. *Edward A. Lincoln*. 284

RESULTS OF REPEATED MEASUREMENT OF PUPIL ACHIEVEMENT. *Gertrude Hildreth*. 286

AN EXPERIMENT ON THE RELATIVE EFFICIENCY OF STUDYING FRENCH VOCABULARY IN ASSOCIATED PAIRS VERSUS STUDYING FRENCH VOCABULARY IN CONTEXT. *Louise C. Seibert*. . 297

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION 315

Published Monthly Except June to August by

Warwick & York, Inc.,

10 East Centre Street Baltimore, Md.

Entered as Second Class matter November 15, 1921, at the Post Office at Baltimore, Md., under the Act of March 3, 1879; additional entry as second class matter at York, Pa.

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

ESTABLISHED 1910

Editor: HAROLD RUGG

Board of Associate Editors

J. CARLETON BELL — ARTHUR I. GATES — RUDOLF PINTNER — E. L. THORNDIKE
FRANK N. FREEMAN — V. A. C. HENMON — L. M. TERMAN — GOODWIN B. WATSON

Assistant to the Editor: FRANCES M. FOSTER

The Journal is published monthly, except June to August, by Warwick and York, Inc., Baltimore, Md., and York, Pa.

The price is $5.00 a year, payable in advance; foreign postage 30 cents extra. Single current issues are 70 cents each, and subscriptions for less than a year are charged at single issue rate for each number ordered. Back volumes are $6.00 each; single back numbers, 90 cents each, except, when more than three years old, the price is $1.00 each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of the next succeeding number.

Manuscripts for publication, books or other materials for review, and correspondence concerning editorial matters should be addressed to Harold Rugg, Editor, 425 West 123rd Street, New York City. Subscriptions should be sent to the publishers.

WARWICK AND YORK, INC., *Baltimore, Maryland*

INTRODUCTORY PSYCHOLOGY FOR TEACHERS

By EDWARD K. STRONG, JR.

$2.20 + 12¢ postage

WARWICK & YORK, INC., *Baltimore, Maryland*

TEACHERS' HEADQUARTERS

European

Plan

Entirely

Fire-Proof

Centrally

Located

EDWARD DAVIS

Manager

HOTEL RENNERT

BALTIMORE, MARYLAND

SCHOOL AND SOCIETY

A weekly journal covering the field of education in relation to the problems of American democracy.

The objects of SCHOOL AND SOCIETY are the advancement of education as a science and the adjustment of our lower and higher schools to the needs of modern life.

It encourages discussion of all problems from every point of view; it pays special attention to foreign affairs; it publishes accounts of educational events of the past week and research on fundamental problems.

Subscription $5.00—Single Numbers 15 Cents

[*SCHOOL AND SOCIETY will be sent to new subscribers for thirteen weeks for one dollar, which may be sent at the risk of the publisher.*]

THE SCIENCE PRESS

Grand Central Terminal - - New York, N. Y.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXI April, 1930 Number 4

SEX DIFFERENCES: INTEREST IN TASKS REQUIRING MECHANICAL ABILITY AND MOTOR SKILL

HARVEY C. LEHMAN

Ohio University

AND

PAUL A. WITTY

University of Kansas

Popular thinking assumes conspicuous differences between the sexes in most traits. The problem of sex differences in abstract intelligence has received the attention of numerous psychologists. Terman, Hollingworth, Whipple and others have presented and discussed sex differences in this regard.[1] Bibliographies on the subject of sex differences have been presented by Winsor and Allen.[2] The problem of sex differences apparently is destined to receive an increasing amount of attention.

This paper will present certain salient sex differences in participation in certain activities which appear to require manual dexterity and motor skill. Although participation in these activities may not reveal motor ability, it undoubtedly betokens interest. And insofar as interest may be considered symptomatic of ability, participation may be considered an indicator of mechanical ability and skill. Mechanical ability and motor skill are general terms and it is obvious that there are many gradations and kinds of such abilities. In this paper the writers will report the extent to which unselected children engage in certain activities usually alleged to require motor ability or mechanical skill.

METHOD

Over five thousand children were asked to check from a comprehensive and catholic list of two hundred play activities only those

activities in which they had voluntarily engaged during the preceding week. The directions were specific, the teachers were instructed in giving the tests, and the data were obtained from unselected school children in order that the results of the investigation might be reliable. Data were obtained from children in the following Kansas towns: Kansas City, Lawrence, Bonner Spring, and Moran. The older individuals were for the most part students of the University of Kansas.

In order that seasonal differences might be taken into account the list was checked by the above groups on each of three different dates, Nov. 7, 1923, Feb. 20, 1924, and Apr. 30, 1924. The same list of activities was checked on each of these dates and the same procedure in administering the test was used in each testing. The total number of individuals from whom data were obtained is indicated in Table I.

TABLE I.—NUMBER OF INDIVIDUALS INCLUDED IN THREE INVESTIGATIONS OF PLAY BEHAVIOR

Ages	Boys			Girls		
	Nov., 1923	Feb., 1924	Apr., 1924	Nov., 1923	Feb., 1924	Apr., 1924
8½	98	90	80	100	97	99
9½	169	161	144	174	139	144
10½	182	169	160	215	199	176
11½	187	167	184	235	222	220
12½	249	201	176	326	289	266
13½	280	231	259	269	235	263
14½	274	252	238	301	282	278
15½	230	247	247	261	244	256
16½	210	181	193	251	223	235
17½	145	130	146	182	208	193
18½	115	170	130	120	174	167
19½	95	119	102	101	115	93
20½	50	73	59	73	124	76
21½	53	57	43	41	85	66
22 and above	79	105	68	44	114	88

RESULTS

Data were assembled according to age and sex. Tables and charts were prepared showing the per cents of individuals in each age group

who had participated in each of the two hundred activities of the Play Quiz during the course of one week. The amount of sex difference may be studied best by examination of the graphs made for each of the two hundred activities. Lack of space prevents the inclusion of these graphs and the data from which they were made. However, a survey of the sex differences may be obtained by examining the lists of the activities in which the largest sex differences were found.

TABLE II.—PLAY ACTIVITIES MORE COMMONLY PARTICIPATED IN BY BOYS THAN BY GIRLS AT PRACTICALLY EVERY AGE LEVEL

*1. Football
*2. Using a hammer, saw, nails, etc. for fun
*3. Wrestling
*4. Marbles
*5. Riding a bicycle
*6. Climbing porches, trees, fences, posts, etc.
*7. Throwing rocks and stones
8. Playing cowboy
*9. Boxing
10. Whistling
*11. Shooting a gun
12. Playing Indian
13. Playing robber and police
*14. Basket ball
*15. Baseball with a hard ball
*16. Driving an auto
17. Watching athletic sports
18. Smoking
*19. Jumping for distance
*20. Jumping for height
*21. Mumbly peg
*22. Flying kites
*23. Making or using a wireless or other electrical apparatus
*24. Snowball fights
*25. Pole vaulting
*26. Just playing catch
*27. Coasting on a wagon
*28. Rolling an auto tire
*29. Fishing
*30. Playing pool
*31. Pitching horse shoes
*32. Playing with bows and arrows
*33. Swimming
*34. Digging caves or dens
*35. Spinning tops
36. Playing with toy trains, ships, autos, wagons, etc.
37. Playing fire engine (or hook and ladder)
*38. Running races
39. Matching pennies
*40. Coasting on a coaster
41. Playing bandit
42. Playing soldier
43. Playing with pet dogs
*44. Playing with an indoor or playground ball
*45. Horseback riding
*46. Rolling a hoop
*47. Walking on stilts
*48. Hunting
49. Building or watching bonfires
*50. Building a dam

Table II presents the activities which were found to be much more commonly participated in by boys than by girls. The activities of Table II are arranged in order of merit, those showing the largest amount of sex difference being listed toward the top of the table. Table III presents a similar list showing the play activities in which the girls participated much more frequently than the boys. At the

time the data were compiled for Tables II and III the writers were engaged in writing a book on the psychology of play activities.[3] At that time this report was not contemplated. The lists of activities presented in Tables II and III were compiled therefore without regard for sex differences in mechanical skill and motor ability.

The writers have attempted to identify in Tables II and III those activities most likely to involve an appreciable amount of mechanical

TABLE III.—PLAY ACTIVITIES MORE COMMONLY PARTICIPATED IN BY GIRLS THAN BY BOYS AT PRACTICALLY EVERY AGE LEVEL

1. Playing with dolls
2. Visiting or entertaining company
3. Playing house
4. Just singing
*5. Sewing, knitting, crocheting
*6. Playing the piano (for fun)
*7. Writing letters
*8. Jumping or skipping rope
9. Dressing up in older folks' clothing
10. Playing school
*11. Social dancing
*12. Jacks
13. Playing movie actor or actress
14. London bridge
*15. Stringing beads
16. Going to parties or picnics
*17. Folk-dancing
18. Just imagining things
19. Looking at pictures
20. Gathering flowers
21. Going to entertainments, concerts, and so on
22. Playing with pet kittens
23. Playing Sunday School
*24. Drop the handkerchief
*25. Cutting paper things with scissors
26. Teasing somebody
27. Just hiking or strolling
*28. Hop, skip, and jump
29. Playing nurse
30. Listening to the victrola
31. Sleigh-riding
32. Reading short stories
33. Telling or guessing riddles
34. Telling fortunes or having fortunes told
*35. Hop-scotch
36. Hide-and-seek
37. Playing store
38. Playing other make-believe games
39. Statuary
40. "Here I come" "Where from?"
41. Tin-tin
42. Other singing games than London bridge
43. Old witch
44. Pussy wants a corner
*45. Making mud pies, mud dolls, etc.
*46. Painting with water-colors
*47. Making a scrapbook
48. Taking snapshots
*49. Roller-skating

skill or motor ability. It is of course true that almost any activity involves some degree of motor ability. However, it is apparent that some activities require much more mechanical ability than others. When generous allowance is made for erroneous judgment in the above regard, it is evident that the boys participated much more frequently than the girls in activities involving these abilities. The activities thought by the writers *most likely* to involve an appreciable amount

of mechanical ability or a high degree of motor skill are marked with asterisks (*).

Table III contains only sixteen such activities and Table II contains thirty-seven. Since all of the activities were voluntarily participated in, it seems logical to infer that boys obtain greater pleasure than do girls from participation in activities which require a high degree of mechanical ability or motor skill. It might be assumed also that the boys' greater pleasure is due in part to their greater success in mechanical and motor pursuits.

Of course various interpretations might be placed upon these results. It is obvious that sex differences in background of experience are in themselves very marked and perhaps sufficient to account for sex differences in the types of performances herein discussed. It is conceivable that a list of motor activities might be assembled which would reveal more frequent participation among girls than among boys. The list used was assembled *upon the basis of the children's own reports of their play.* The list represents, therefore, the child's own voluntary and characteristic play life. It is significant that responses to such a list show marked sex differences in frequency of participation in mechanical or motor activities. Some of the activities are characteristically boys' games; others are considered girls' play.

It seems obvious that mechanical ability and motor skill are not unitary traits but generic terms which include a multitude of specific functions or abilities. This hypothesis is supported by the findings of three different investigators; namely, Perrin,[4] Muscio,[5] and Barton.[6] These workers have found that the relatively simple motor abilities correlate with each other only slightly. The question arises therefore whether the so-called mechanical and ability tests measure anything more than *acquired abilities.** In other words, do these tests measure innate ability or do they test merely specific acquired modes of response? If the tests of mechanical ability are merely measures of previous experience and habit acquisition, one is not justified in designating them motor or mechanical ability tests unless he grants that the ability is an acquisition and not an innate predisposition. For example, it is apparent that girls are more skillful than boys in sewing, knitting, crochetting, etc. But does this mean that men *cannot* become as competent as women in these tasks with *equal practice and motivation?* Measurement of sex differences in the above "abilities"

* The Stenquist Mechanical Aptitude Test, the MacQuarrie Test for Mechanical Ability, and others.

would not be measurement of innate ability but merely a statement of present attainment.

It is possible to place numerous interpretations upon the differences cited in this paper. Some might assert that the sex differences in interests are due entirely to differences in the experiences of the sexes, resulting largely from differences in social customs, mores and taboos. On the other hand some might contend that the differences are explained easiest by the hypothesis that there are minor differences in certain aspects of physical growth and development which effect differences in behavior. Some might insist that boys participate more frequently than girls in motor activities because it is easier for boys to attain success in these activities; that boys possess by original nature physical attributes that make them acquire motor skill more easily than girls.

Regarding the origin of the sex differences displayed in Tables II and III there is little to be said with finality.

> There is, unfortunately but little in the data at hand which serves to indicate whether the popular opinion as to masculine superiority in tasks of this (mechanical) sort would be justified if the material used were equally familiar to both sexes. That boys and men on the average do greatly exceed women and girls in the ability to manipulate mechanical contrivances is not only a matter of common observation but has been brought out very clearly in an extensive investigation on mechanical abilities which has been carried out at the University of Minnesota, the results of which will be published in the near future. Woolley reports that among her working children the only major sex differences shown by any of the tests used was a great superiority of the boys in opening the puzzle box.[7] (Pp. 453–454.)

The writers are concerned only with showing differences in the experiential backgrounds of the sexes. The data presented by them give clear evidence that boys participate in certain motor activities more frequently than girls. The data suggest that marked sex differences in voluntary participation in various activities may be assumed logically to be indicative of interest in these activities. The nature of the motor activities suggests that the varying experiences of the sexes are potent forces in effecting the differences. These experiences may be traceable to social customs, attitudes, and values. The latter fact seems so conspicuous as to suggest that alleged innate differences in motor ability previously reported may be due to differences in attitude and experience, the product almost solely of environmental factors. It may be that differences in performance

in the so-called mechanical ability tests are therefore not tests of *ability* at all but merely tests of the present level of performance. Perhaps certain fundamental assumptions of those who presume to test these "abilities" will need to be revised or modified.

References

1. Terman, L. M.: "The Measurement of Intelligence." Boston: Houghton Mifflin Co., 1916. Pp. XVIII + 362. (Pp. 68ff.)
   Hollingworth, L. S.: Variability as Related to Sex Differences in Achievement. *American Journal of Sociology*, 1914, Vol. XIX, pp. 510–530.
   Whipple, G. M.: Sex Differences in Intelligence Test Scores in the Elementary School. *Journal of Educational Research*, 1927, Vol. XVI, pp. 111–118.
   Whipple, G. M.: Sex Differences in Army Alpha Test Scores in the Secondary School. *Journal of Educational Research*, 1927, Vol. XVI, pp. 269–276.
2. Winsor, A. L.: The Relative Variability of Boys and Girls. *Journal of Educational Psychology*, 1927, Vol. XVIII, pp. 327–336.
   Allen, C. N.: Studies in Sex Differences. *Psychology Bulletin*, 1927, Vol. XXIV, pp. 294–304.
3. Lehman, H. C. and Witty, P. A.: "The Psychology of Play Activities." New York: A. S. Barnes and Co., 1927. Pp. XVIII + 242. (Pp. 93ff.)
4. Perrin, F. A. C.: An Experimental Study of Motor Ability. *Journal of Experimental Psychology*, 1921, Vol. IV, pp. 24–56.
5. Muscio, B.: Motor Capacity with Special Reference to Vocational Guidance. *British Journal of Psychology*, 1922, Vol. XIII, pp. 157–184.
6. Barton, E.: "Correlations among Motor Abilities." A. B. Thesis, 1926. (Filed in University of Wisconsin Library. Quoted by C. L. Hull in "Aptitude Testing." World Book Co., 1928, p. 21.)
7. Goodenough, F. L.: The Consistency of Sex Differences in Mental Traits at Various Ages. *Psychology Review*, 1927, Vol. XXXIV, pp. 440–462.

A STUDY OF MIND-SET IN RURAL AND CITY SCHOOL CHILDREN[1]

J. M. SHALES

Associate Professor of Rural Education, Ball State Teachers College, Muncie, Indiana

It is generally accepted psychological principle that the reaction of a human being to any particular situation is to a large extent determined by his mind-set or attitudes. This is true not only for conditions which immediately precede the reaction, but also for the total of his past experience. These more permanent dispositions, which are the infinitely integrated summation of all interactions of a personality with its environment, are very potent determiners of response to fresh stimuli. They form a river of tendencies to react, which gains definiteness of direction the farther it proceeds from its source in original nature. New situations with their corresponding reactions, like tributaries of various size and character, produce their effects on this river, and are in turn largely directed by it. Some new set of circumstances may assume cataclysmic proportions in the life of an individual, forming a dam which almost entirely diverts the stream of dispositions.

In view of this principle, it would be reasonable to expect that children reared in the open country would have what might be termed a "rural" mind-set, and that such attitudes would give a characteristic color to their reactions to situations. Similarly, we would expect the reaction of children raised in the city to be in part determined by the urban quality of their experience. The present study is an attempt to find out to what extent these expectations are justified.

The following twenty questions, each so designed as to permit intelligent answers in terms of either country or city experience, were submitted to groups of country and city children:

1. Where do you get flowers for your table?
2. What is in a block?
3. Where can children play under the trees?

[1] This study was carried out under the direction of Dr. Paul J. Kruse of the Department of Rural Education at Cornell University. The writer is also indebted to Dr. E. Laurence Palmer of the same department and to the many rural teachers throuthout the State of New York, whose cooperation made possible the collection of data from rural school children. The assistance of Miss Cornelia B. Roach and Miss Ada James of the Buffalo City Schools in securing replies from city children is also gratefully acknowledged.

4. What are pumps for?
5. What is often seen near a bank?
6. Can you buy cloth at the store where you get groceries?
7. Where do you get the milk you drink?
8. What are bulbs used for?
9. Have you ever helped your father all day at his regular work?
10. Where do you get the water you drink?
11. At what time do grown-ups go to bed?
12. Do men work as many hours a day in winter as in summer?
13. If you had no watch and could not ask another person, how could you tell the time?
14. What does a floor-walker do?
15. What is the first thing you would do in case of fire?
16. What are rails for?
17. How many rooms are there in a school-house?
18. What is a curb?
19. What kind of place is good for swimming?
20. Is there a barn near most houses?

The questions were submitted to two groups of elementary school pupils from the fourth to the eighth grade inclusive. One group comprised 464 children attending a typical school in the city of Buffalo. Questions were presented on a mimeographed sheet with spaces for answers below each question. The other group consisted of the same number of unselected rural children from the same grade range. The questions were published in the Cornell Rural School Leaflet for September, 1927, and sent to every rural school in New York State. Instructions were given to write the question numbers, followed by their respective answers, on a sheet of foolscap paper and send them to the writer. Of the several thousand answer sheets received, the first 464 to complete a grade distribution similar to the city group were used for the study. A few answer sheets which were very incomplete or illegibly written were thrown out. Answers from two schools with sixteen and eighteen rooms respectively and located in fairly large towns were rejected in order to obtain a more purely rural group. Both groups were asked to write their ages and school grades at the top of their answer sheets. Distribution by grades for each group was as follows:

Grade	Number
IV	100
V	104
VI	77
VII	91
VIII	92
Total	464

One hundred answer sheets from each group were first analyzed, and a set of criteria for judging the answers was worked out. These criteria were followed rigidly in scoring all the answers. Each reply was marked with one of the symbols, *R*, *C*, *E*, or *O*, to indicate the set of mind judged to have prompted the answer. *R* represented a rural set; *C*, a city set; *E*, an answer that might have originated from either; and *O*, an entirely irrelevant answer or a question omitted.

Tabulation of answers marked according to the above scheme yielded results as shown in Table I. The rural group gave answers which were predominantly indicative of a rural mind-set, except in questions 8, 11, 14, 15, and 18. In question No. 8 What are bulbs for? the answer "Electric lights" was considered a city answer, and "for planting for flowers" a country answer. The small reverse difference

TABLE I.—DISTRIBUTION OF ANSWERS ACCORDING TO TYPE FOR RURAL AND CITY GROUPS, STATED IN PER CENT OF TOTAL ANSWERS

Question	Rural group				City group			
	R	*C*	*E*	*O*	*R*	*C*	*E*	*O*
1	89.44	6.89	3.01	.64	79.31	15.94	3.12	1.50
2	44.82	38.36	5.17	11.63	19.18	51.72	6.89	22.19
3	41.59	15.08	31.25	12.06	11.63	41.37	28.23	18.75
4	90.51	1.29	6.68	1.50	60.56	10.56	16.59	12.28
5	78.66	17.02	0.00	4.31	25.00	62.50	3.01	9.48
6	54.52	45.04	0.00	.43	3.23	93.53	1.07	2.15
7	88.14	11.20	.21	.43	48.27	50.43	.21	.43
8	41.81	50.86	4.09	3.23	23.70	64.43	.21	11.63
9	50.21	48.06	.43	1.29	22.41	75.43	.43	1.72
10	94.39	3.12	1.07	1.29	4.09	90.94	.43	4.52
11	37.71	43.10	7.32	11.85	5.17	67.88	7.32	19.61
12	80.81	16.59	0.00	2.58	42.24	50.43	1.50	5.81
13	85.99	4.09	.64	9.26	65.30	14.65	3.66	16.37
14	11.20	50.21	11.63	6.93	1.07	63.57	13.79	21.55
15	37.50	42.02	18.31	2.15	.64	63.79	31.46	4.09
16	46.76	45.90	4.74	2.58	1.07	81.68	5.38	11.85
17	80.38	13.57	0.00	6.03	16.81	62.71	.64	19.82
18	15.73	50.43	4.09	29.74	.43	66.37	9.05	24.13
19	27.37	23.48	46.55	2.58	1.50	58.62	31.89	7.97
20	82.97	16.37	0.00	.64	15.08	81.25	0.00	3.66
Average	59.02	27.13	7.25	6.55	22.33	58.39	8.24	10.97
	$N = 464$				$N = 464$			

in the answer totals probably indicates that electric lighting is becoming rather common in rural areas, and that city children are fairly familiar with the use of bulbs to grow certain flowers.

In question No. 11 (At what time do grown-ups go to bed?) the hour of 9:30 was arbitrarily set as the dividing line between rural and city experience. It is strongly suspected that many of the answers of the younger children were valiant guesses rather than facts, and it is quite possible that the point of division was set too low. In question No. 14, (What does a floor-walker do?) a decided majority of the rural children knew the store and factory functions of floorwalkers, which fact is taken as evidence of city experience. Question No. 15 regarding conduct in case of fire showed a small difference in favor of urban reactions by rural children. All answers presupposing the existence of an available fire department which might be called upon for help, were taken as an indication of city experience. The writer is at a loss to explain the apparent reliance upon such aid by rural children unless the glamor and romance of city fire fighting methods, as observed during visits to a city, had carried away their imagination. The answers from many of the pupils from one-room schools would lead one to think that "the woods were full of" fire-engines. Question No. 18 (What is a curb?) was considered to have been answered from city experience when curb was defined in relation to a street. Such definition admitted acquaintance with village streets, and therefore, no doubt, included many who were not consciously referring to city situations. Moreover, many answers smacked strongly of quotations from a dictionary, and so did not give a reliable picture of the children's mind-sets.

The city group gave city answers even more consistently than the rural group gave rural answers. Notable exceptions are seen in questions 1, 4, and 13. In the case of the first question, no satisfactory definition to differentiate city from country answers could be made. It was finally decided that getting flowers from stores or florists would be taken to indicate city experience, while gardens, fields and woods would constitute evidence of rural attitude. The great majority of city children gave the answer, "From the garden," which had to be classified as a rural reaction. There is no way of telling whether practically all the homes represented by this Buffalo school have flower-gardens, or whether the pupils are rural-minded enough to know where flowers are generally produced. Question No. 4 (What are pumps for?) was also difficult to give a differentiating definition. The predominant

answer was "To get water from," which was taken as representing a typically rural experience. It is possible that this may, after all, give a fairly good indication of mind-set, since most children who have ever visited the country would probable retain fairly definite memories of having their thirst quenched from the pump in the yard. Question 13 was expected to elicit from the city children an allusion to the numerous clocks to be seen in cities, and from country children evidence of reliance upon the position of the sun to tell the time. On the contrary, sixty-five per cent of the city children were conversant with the method of estimating the time by the sun, and were ready to use it in an emergency. In this respect we may conclude that these groups of rural and city children are similarly minded.

Viewing the results as a whole, however, answers to this set of questions show a decided rural and equally pronounced city mind-set in the respective rural and city groups. Rural children gave 5478

TABLE II.—RANKING OF VARIOUS QUESTIONS ACCORDING TO THE EXTENT OF RURAL AND CITY MIND-SET SHOWN

Rank number	Rural group		City group	
	Per cent rural answers	Question number	Question number	Per cent city answers
1	94.39	10	6	93.53
2	90.51	4	10	90.94
3	89.51	1	16	81.68
4	88.14	7	20	81.25
5	85.99	13	9	75.43
6	82.97	20	11	67.88
7	80.81	12	18	66.37
8	80.38	17	8	64.43
9	78.66	5	15	63.79
10	54.52	6	14	63.57
11	50.21	9	17	62.71
12	46.76	16	5	62.50
13	44.82	2	19	58.62
14	41.81	3	2	51.72
15	41.59	8	7	50.43
16	37.71	11	12	50.43
17	37.50	15	3	41.37
18	27.37	19	1	15.94
19	15.73	18	13	14.65
20	11.20	14	4	10.56

rural and 2517 city answers, while the city group submitted 5419 city and 2073 rural answers. Stated in another way, 59.02 per cent of all answers from the rural group were rural in nature and 27.13 per cent met the criteria of city answers: the city group gave city answers to the extent of 58.39 per cent of their total, and rural answers in only 22.33 per cent of the cases.

To get a clearer picture of the extent to which the various questions revealed rural and city mind-set in the respective groups, the questions were ranked in that order for each group. Table II presents the distribution found.

The average per cent of rural answers from the rural group on the first ten questions in the above ranking was 82.59, while the first ten for the city group brought 74.88 per cent of city answers. This represents a fairly high degree of effectiveness in calling forth these types of responses. The table shows that questions which are effective in showing rural mind-set are not necessarily useful in determining city attitudes, since two questions, No. 10 and No. 20, are the only ones to appear in the first ten with both groups. This would indicate that a large number of questions would give a greater validity in measuring mind-set than a short list as used in this study.

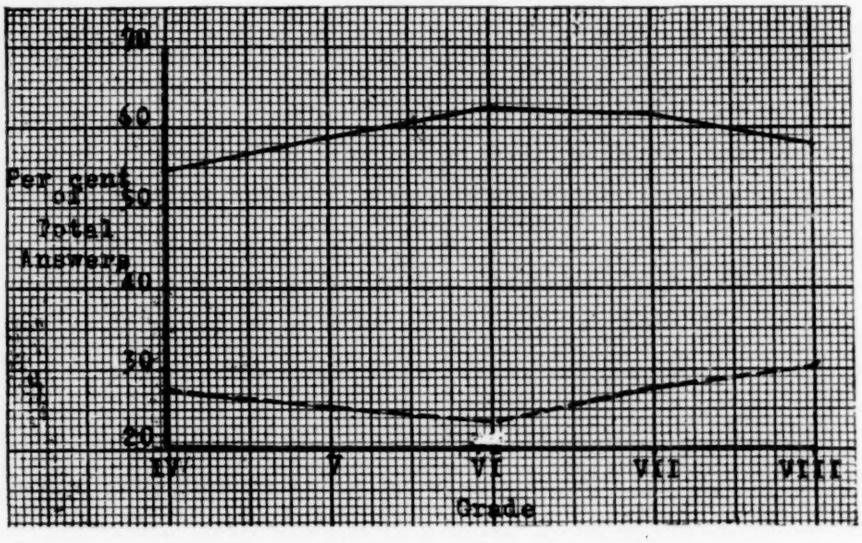
Arrangement of data by school grades gave the distribution of answers shown in Tables III and IV.

TABLE III.—DISTRIBUTION OF RURAL GROUP ANSWERS BY SCHOOL GRADES

Grade		*R*		*C*		*E*		*O*	
	Number of pupils	Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent
IV	100	1089	54.45	547	27.35	140	7.00	224	11.20
V	104	1219	58.60	522	25.09	174	8.36	165	7.93
VI	77	957	62.14	358	23.24	143	9.28	82	5.32
VII	91	1119	61.48	505	27.74	127	6.97	69	3.79
VIII	92	1066	57.93	563	30.59	141	7.66	70	3.80

Table III shows that in rural children there is a rather marked tendency for the proportion of rural answers to increase through Grades IV, V, and VI, and then to decrease in the next two grades, showing a drop of almost five per cent from the high point in Grade VI to

Grade VIII. An exactly parallel variation is seen in the city answers made by this group, as they decrease steadily in the first three grades and then increase rather sharply through Grades VII and VIII. It



Rural answers—continuous line
City answers—broken line

FIG. 1.—Distribution in per cents of rural group answers by grades.

will be noted from the table that the proportion of "*E*," or noncommittal answers remained fairly constant, while the "*O*," or unsatisfactory and omitted replies, show a natural decrease as the child

TABLE IV.—DISTRIBUTION OF CITY GROUP ANSWERS BY SCHOOL GRADES

Grade	Number of pupils	*R*		*C*		*E*		*O*	
		Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent
IV	100	461	23.05	1048	52.40	181	9.05	310	15.50
V	104	437	21.00	1156	55.57	167	8.02	320	15.38
VI	77	342	22.20	950	61.68	128	8.31	120	7.79
VII	91	458	25.16	1105	60.71	146	8.02	111	6.09
VIII	92	375	20.38	1160	63.04	144	7.82	161	8.75

proceeds in his education. The trends of answers are shown graphically in Fig. 1.

The most probable explanation of the peculiar reversal of tendency in the type of answer in Grades VII and VIII is that children are

reaching adolescence about this time. Since this period is credited with bringing a broadening of interests and sympathies and a generally more cosmopolitan view-point, it is reasonable to expect that it may help to make rural mind-set less pronounced.

Table IV shows that city children gave a steady increase in proportion of city answers as they advanced in school. There is a break of one per cent away from this tendency between Grades VI and VII

Rural answers—continuous line
City answers—broken line

FIG. 2.—Distribution in per cents of city group answers by grades.

and an increase of three per cent in rural answers, but the previous tendency is very decidedly resumed in Grade VIII. There is a small variation in the per cent of rural answers given by the various grades and the number of "*E*" answers remains practically constant.

The generally steady decrease in the number of unacceptable answers appears to have resulted in pure gains for the city type. The tendencies to variation through the grades is shown in Fig. 2.

The data assembled in this study would justify the conclusion that the mind-set of this city group becomes more and more urban as pupils advance in school. If the coming of adolescence is a factor in broadening the outlook of country children, we might assume that this period is less of a crisis and produces less pronounced changes in city children. Some justification for such an assumption might be found in the well known fact that city children do not, as a rule, take much part in occupational pursuits until adolescence is at least well advanced, while most farm children do assume many responsibilities.

In other words, school remains the paramount concern of the city child during this period, while it tends to be displaced in the interests of country boys and girls. Another reason which might be suggested for the failure of city children to enlarge their rural experiences is to be found in the traditional attitude of disparagement with which city people look upon things of the country. This feeling is probably less pronounced than it has been in the past, but some of it still remains. Many city children yet believe that it is beneath their dignity to manifest too intimate acquaintance with rural institutions, while knowledge of city ways is for country children an avidly pursued mark of sophistication.

Although chronological age and school grade follow a generally parallel course, it was thought that some further light might be cast on the question of mind-set by examining the answers from the standpoint of the ages of the children. The data were accordingly arranged in this way and are presented in Tables V and VI.

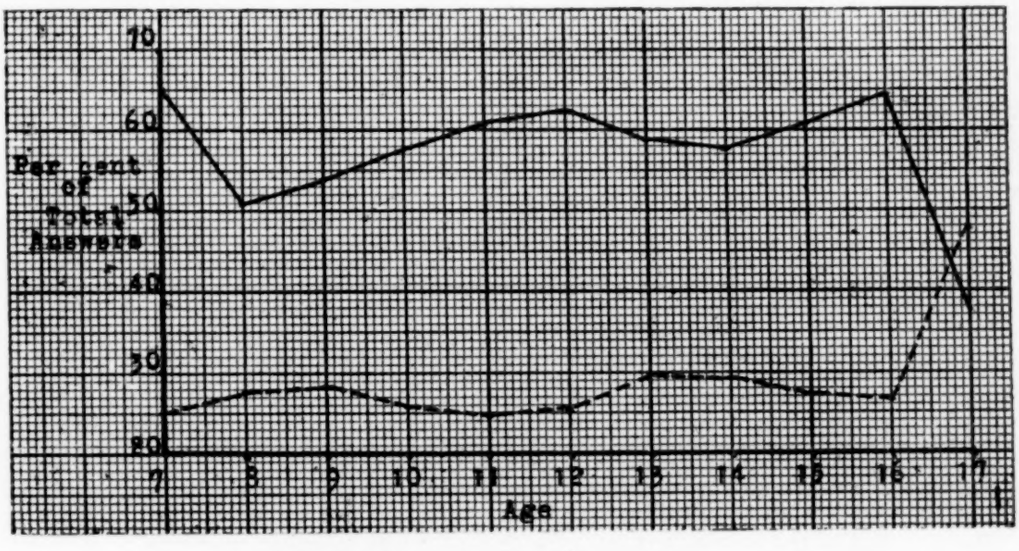
TABLE V.—DISTRIBUTION OF RURAL GROUP ANSWERS BY AGES

Age	Number of pupils	*R*		*C*		*E*		*O*	
		Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent
7	3	39	65.00	15	25.00	2	3.33	4	6.66
8	15	163	50.93	89	27.81	24	7.50	44	13.75
9	40	431	53.87	225	28.12	68	8.50	76	9.50
10	71	818	57.59	367	25.84	118	8.30	117	8.23
11	78	940	60.25	388	24.87	121	7.75	111	7.11
12	82	1020	62.19	413	25.18	120	7.31	87	5.30
13	72	845	58.68	420	29.16	104	7.22	71	4.93
14	58	668	57.58	340	29.31	94	8.10	58	5.00
15	36	435	60.41	196	27.22	59	8.19	30	4.16
16	7	90	64.28	30	21.42	8	5.71	12	8.57
17	2	15	37.50	19	47.50	4	10.00	2	5.00

It will be seen that the rural group shows a wider range of ages by three years than the city group, one year at the lower end of the scale and two years at the upper end. The three seven-year-old children showed a decidedly rural set of mind, while the two seventeen-year pupils exhibited a rather startling turn to city mindedness.

Seven sixteen-year-old pupils appear to follow the tendency of the next lower groups. Since these three ages are not represented in the

city group, and since their numbers are so few, no attempt will be made to explain the apparent vagaries of their answers. Neglecting these three ages, we find the proportion of rural answers from the rural group increasing steadily until age thirteen is reached, when a decrease sets in which amounts to five per cent between the ages of twelve and fourteen. With the age of fifteen the previous tendency to increase appears to be resumed. There is a very slight tendency for the per cent of city answers to decrease as age advances through the twelfth year, after which there is a fairly decided increase for two years. Decrease begins with the fifteenth year and becomes very pronounced with the small group of sixteen-year-old pupils, whose education has been for some reason seriously retarded. The abrupt increase of unsatisfactory answers is further evidence of the retardation of this age group. The variations in proportions of rural and city answers by the rural group are summarized graphically in Fig. 3.



Rural answers—continuous line
City answers—broken line

FIG. 3.—Distribution in per cents of rural group answers by ages.

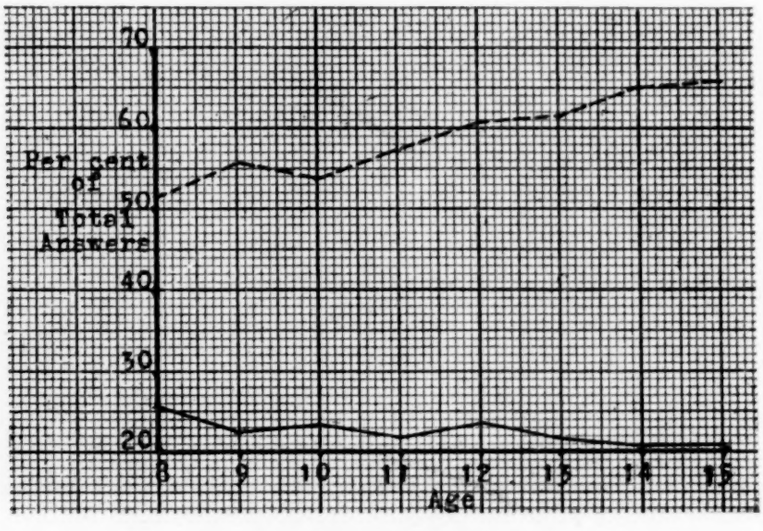
In the city group an age distribution from eight to fifteen years inclusive was found. The per cent of rural answers shows a tendency to decline as age advances, although the total decrease is only five per cent. A slight reverse movement occurs in the age twelve group, but disappears at once. The proportion of city answers grows steadily with advancing years, showing a total increment of almost fourteen per cent. The ten-year-old pupils exhibited a slight, temporary counter-tendency, which was accompanied by a corresponding increase

in the proportion of unacceptable answers. Figure 4 gives a picture of the tendencies in the replies of this group.

TABLE VI.—DISTRIBUTION OF CITY GROUP ANSWERS BY AGES

Age	Number of pupils	*R*		*C*		*E*		*O*	
		Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent
8	16	81	25.31	166	51.87	30	9.37	43	13.43
9	88	388	22.04	985	55.96	159	9.03	228	12.94
10	72	334	23.19	775	53.81	127	8.81	204	14.16
11	84	369	21.96	960	57.14	127	7.55	224	13.33
12	79	374	23.67	957	60.56	131	8.29	118	7.46
13	77	337	21.88	947	61.49	132	8.57	124	8.05
14	33	134	20.30	429	65.00	53	8.03	44	6.66
15	15	61	20.53	196	65.33	15	5.00	28	9.33

At the beginning of adolescence similar fluctuations were seen in the answers of the rural group when the answers were arranged



Rural answers—continuous lide
City answers—broken line

FIG. 4.—Distribution in per cents of city group answers by ages.

according to age as were found when they were assembled by school grade. This gives further confirmation to the theory that this period may account for some broadening of view-point among country children. The arrangement by ages, however, showed a tendency for

this group between the ages of fourteen and sixteen to become even more rural-minded than before. This points to the likelihood of selection coming into play when the children are about fourteen years old. It is very probable that many of the brighter and more urbane pupils withdraw at that age to attend high schools in larger centers, leaving the slower, less adaptable students in the rural, elementary school.[1]

The study of these particular groups of pupils by this technique leads to the following conclusions:

1. That rural children have a pronounced rural set of mind, and that city children have a decided, characteristic urban type of mind-set.

2. That rural children become more rural-minded until about twelve years old, then less so for the next two years, after which the previous trend is resumed. Progress through the grades shows increasing rural mind-set through Grade VI, after which such attitude becomes progressively less pronounced. City children tend to become steadily more city-minded as they grow older and as they proceed in their education.

It is not within the scope of this study to develop the educational bearings of the above conclusions. One implication for teaching method is, however, so apparent that it cannot well be neglected. If country and city children possess stores of percepts which are largely of different types, it very evidently follows that teaching should be based upon, and proceed in terms of, such characteristic experiences. It is not intended to argue that school procedure should aim to augment one type of meanings at the expense of the other, but simply that a child's present equipment of percepts should be the point of departure, if teaching is to be as effective as it should. This principle is as old as teaching itself, but teachers and textbooks have been, and are still, far from giving it the fullest practical recognition. In the light of the above findings, a country teacher should make the rural-mindness of her pupils the foundation of her teaching technique, while the urban-mindness of city children should serve a like function for the city teacher.

The present study might well be extended to trace tendencies in pupils of later ages and in the high school grades. It should also be interesting to trace the relations between mind-set and the following factors among others:

[1] Compare *Texas Survey Report*, Vol. IV, Chap. IV.

(*a*) Intelligence as expressed by IQ.
(*b*) Sex of pupils.
(*c*) Foreign and native-born pupils.
(*d*) Stability of residence in country or city.
(*e*) Amount of reading done by the children.
(*f*) Education of parents.
(*g*) Occupations of parents.

AN EXPERIMENTAL COMPARISON OF THE MOVEMENTS IN MANUSCRIPT WRITING AND CURSIVE WRITING

WILLIAM HENRY GRAY

Kansas State Teachers College

It was the purpose of this investigation to analyze and compare the movements executed by writers who are trained in the newly-introduced, so-called manuscript writing with those executed by persons who are trained in the ordinary or cursive writing. The investigation determined with precision the hand and arm positions at various points in writing: at the beginning and end of a line of writing, at the beginning and end of a word, and at the top and bottom of a letter stroke; and the time consumed in passing from one position to the other. It also determined the various elements of each movement by studying separately the movements of the hand, arm, and those of the fingers. Internal speed changes within the two types of writing were minutely analyzed and compared.

The subjects used in the study included ten adults and twenty children. The children were chosen from the public school system of Winnetka, Illinois, and from the Elementary School of the University of Chicago. The children from Winnetka had been trained in manuscript writing, while those from the University were trained in cursive writing. All children were chosen from Grade V on the basis of tests and were selected because they were shown by the tests to be writers of average speed and quality. The adults were obtained among students of the University of Chicago. All of them wrote the two styles of writing. Accordingly records of the two types were taken from each individual, making possible individual as well as group comparisons.

The records were made by photographing the movements of the hand and arm by means of a kinetoscopic camera. The camera was mounted directly above the hand of the writer and could be speeded up to thirty exposures per second. The exposures were controlled as to rate so that it was possible to determine with precision the time intervals between successive positions of the pen.

The record of a single sentence was taken in each case. The sentence written by the adults was, "*The job requires extra pluck and*

zeal from every young wage earner"; that written by the children, *"We use writing in school subjects."*

The method of analyzing the records was as follows:

The films were first run through a projecting machine to obtain a general notion of the main features of the writing movements. As a second step, drawings were made of the hand positions at the beginning and end of certain lines of writing, the beginning and end of a specially selected word, and the top and bottom of a specially selected letter. As a third step, the writing was traced and the position of the pen point at each successive exposure of the camera was indicated upon this tracing of the writing. The plotting of pen positions made it possible to determine the distance travelled by the pen in successive units of time. A statistical comparison was made of the time intervals consumed in the two types of writing.

Special graphical analysis was made of the movements made in writing the word *pluck* in the case of adults, and of the word *school* in

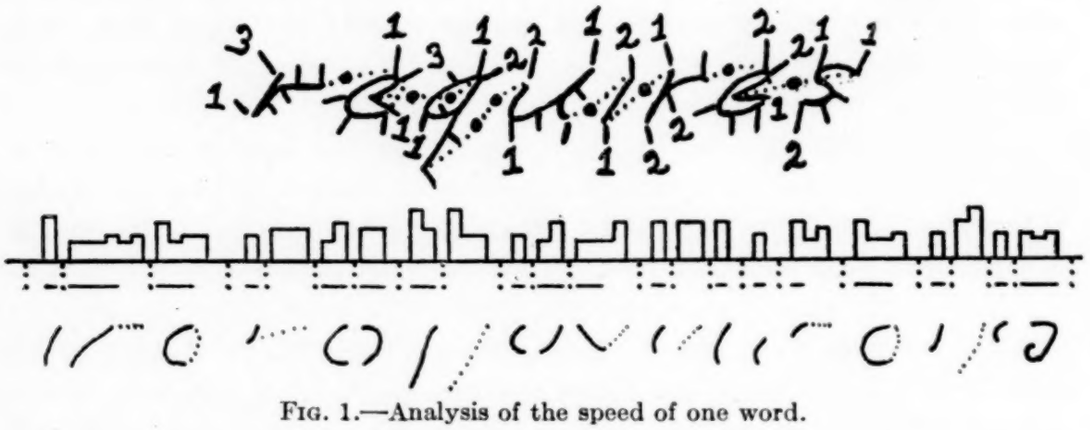


FIG. 1.—Analysis of the speed of one word.

the case of children. Any other words would have served the purpose equally well but these words were chosen as typical because the movements were well under way before the words in question were written.

Figure 1 is an illustration of the plotting of the word *requires*, taken from the record of the manuscript writing of one of the subjects, together with the speed chart constructed from the plotted word. A crowding of the points on the writing means a slowing down of the movement, while a spreading of the points indicates an acceleration. The numbers indicate pauses. The dotted lines represent movements which were made with the pen raised from the paper, the position of the pen at the end of each exposure being represented by larger dots.

In the chart at the bottom of Figure 1 each horizontal unit represents one exposure of the camera or one twenty-second of a second. The height of the column above each unit represents the distance in fifths of a centimeter which the pen travelled during the corresponding exposure of the camera. The absence of a column indicates a pause of one twenty-second of a second. For example, three exposures of the camera occurred after the initial one at the beginning of the word *requires*. This pause is indicated in the chart by the horizontal line at the beginning, three units long. During the next exposure of the camera, the pen travelled from the top to the bottom of the letter *r*. This is indicated by a column three and one-half units tall. In transferring the length of strokes from the enlarged writing to the squared paper, a pair of dividers were used. A series of columns represents a movement of the pen. The speed of the movement is represented by the height of the successive columns and the duration of the movement by the number of columns. Pauses in the movement are indicated by three dots placed vertically below the curve. Movements of the pen made in contact with the paper are distinguished from those made in the air by horizontal lines drawn beneath the curve. The spaces between these horizontal lines represent the movements of the pen when raised from the paper. The part of the letter which is made during each of the units of movement is also traced underneath the curve, movements off the paper being indicated by dotted lines.

Representative drawings of hand positions are shown in Figures 2 and 3 together with reproductions of parts of the two styles of writing produced under the camera by the individual in question. Where it was necessary to superimpose one drawing upon another, a dotted line is used to represent the second position of the hand. Figure 2 shows hand positions for the manuscript writing, while Figure 3 shows hand positions for the cursive writing.

If the positions at the beginning of the word *requires* are examined first, it will be noted that for the two types of writing there is great similarity. The forearm is held almost perpendicular to the base line of writing. The hand is turned slightly to the right. There is a difference, however, in the angle at which the pen is held and the hand is turned over somewhat more in the manuscript writing. In the cursive writing the penholder points almost in the direction of the forearm, while in the manuscript writing it is inclined towards the right at an angle of about forty-five degrees. In both styles of writing

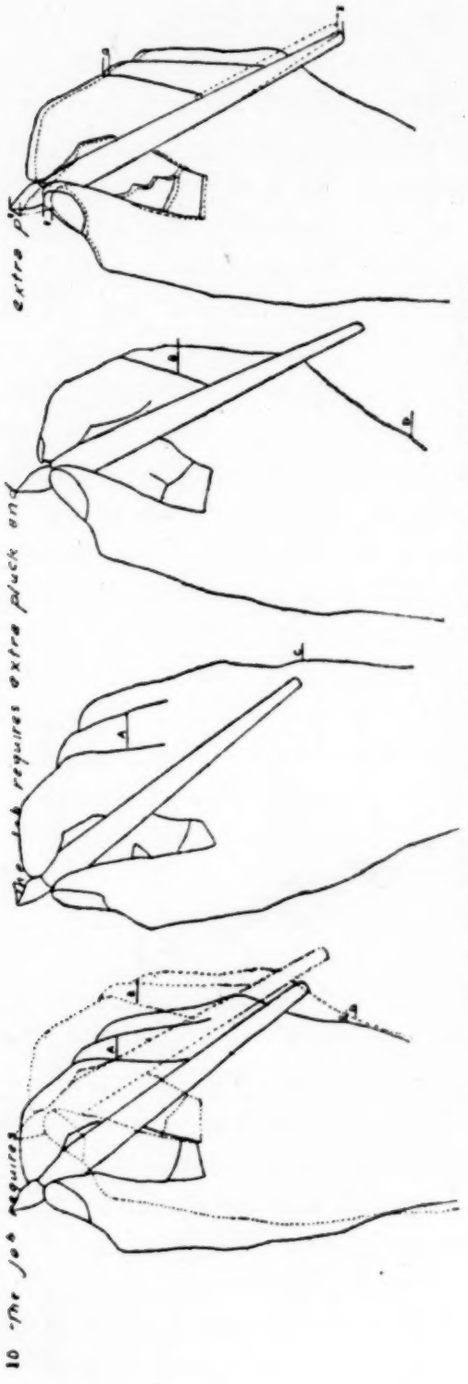


Fig. 2.—Hand positions, manuscript writing of a selected individual.

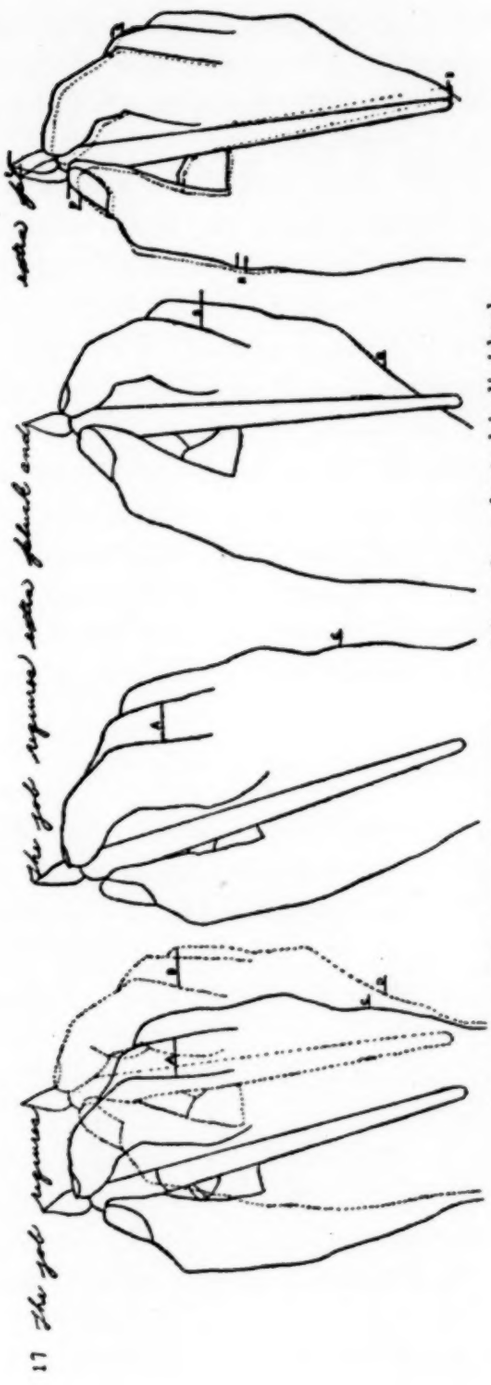


FIG. 3.—Hand positions, cursive writing of a selected individual.

the penholder is grasped lightly by the thumb and the first two fingers. The forefinger rests slightly to the right side of the penholder and the thumb presses against the opposite side at a point slightly farther from the pen point than the finger.

The movement of the hand along the line is also similar for the two types of writing. In both cases the hand turned over towards the right as it progressed along the page. This is evidenced by comparing the lines A and B (Figures 2 and 3) drawn from the edge of the index finger to the edge of the middle finger. These lines are of the same length so that the outlines of the two fingers are shown to be closer together in the hand positions at the ends of the lines of writing. Lines C and D drawn at the wrist joint show that the elbow has been drawn back more in the manuscript writing as the writing progressed.

The movement from the top to the bottom of the letter *l*, as evidenced by the positions of the hand and pen at points E to I in the drawings at the right of Figures 2 and 3, is made exclusively by the fingers and thumb in both the manuscript writing and the cursive writing.

Results of the Investigation

The study of hand positions revealed very little difference for manuscript writing and cursive writing and the discrepancy that does exist appears to be due to individual differences rather than to differences caused by one or the other types of writing. Movement along the line, within words and on the letter strokes is very similar in the two types of writing. For this reason, the presentation of results will be confined to the speed changes which take place within the two types of writing.

Table I indicates the time spent by adults in movement within words. The table is to be read as follows: Subject 1 consumed 504 units of time in movement within words for manuscript writing as compared with 375 units of time for cursive writing making a difference in time of 129 units.

It will be noted that Subject 5 requires less time for the manuscript writing than for the cursive writing. A conversation revealed that this subject uses manuscript writing almost exclusively. Subjects 4 and 9 show a relatively small difference in the time required for each type of writing. The average for the group is 565.4 units of time for manuscript writing and 466.6 units of time for the cursive hand, making a difference of 98.8 units of time or about 4.5 seconds.

A similar comparison is shown for the children in Table II. Since the children from whom the manuscript records were secured form a different group from those from whom the cursive records were

TABLE I.—COMPARISON OF TIME CONSUMED BY ADULTS IN MOVEMENT WITHIN WORDS

Subject	Time in 22nds of a second		Difference in time
	Manuscript	Cursive	
1	504	375	129
2	530	408	122
3	570	413	157
4	527	499	28
5	525	567	− 42
6	655	543	112
7	784	535	249
8	519	433	86
9	411	385	26
10	629	508	121
Average	565.4	466.6	98.8

TABLE II.—COMPARISON OF TIME CONSUMED BY CHILDREN IN MOVEMENT WITHIN WORDS

Manuscript		Cursive	
Subject	Time in 30ths of a second	Subject	Time in 30ths of a second
21	550	31	271
22	497	32	312
23	462	33	401
24	344	34	498
25	348	35	353
26	467	36	382
27	402	37	385
28	385	38	392
29	435	39	408
30	447	40	422
Average	433.7		382.4

obtained, no individual comparison can be made. It will be noted, however that some cursive writers write more slowly than some manuscript writers. On the average, the cursive writers write more rapidly than do the manuscript writers, the average length of time

being 433.7 units for manuscript writing as opposed to 382.4 units for cursive writing. This makes a difference of 51.3 units or about 1.7 seconds.

It will be noted that the difference in time is much greater for adults than for children which indicates a greater increase in speed with age for cursive writing.

A comparison of the time spent by adults in pauses at the end of strokes is given in Table III. This table is to be read as follows: Subject 1 spent 128 twenty-seconds of a second or 25.4 per cent of the time required for manuscript writing in pauses at the end of strokes as compared with 79 twenty-seconds of a second or 21.1 per cent of the time required for cursive writing. Each individual spends a larger amount of time in pauses in the manuscript writing than in the cursive writing. The average for the group is 41.8 per cent for manuscript writing as compared with 32.8 per cent for cursive writing.

The comparison of time spent by children in pauses at the end of strokes is given in Table IV. Subject 21, a manuscript writer spent 409 units or 74.4 per cent of the writing time in pauses; subject 31, a cursive writer, spent 112 units or 41.3 per cent of the writing time in pauses. The average for the group is 65.7 per cent for manuscript writing as compared with 51.9 per cent for cursive writing.

These figures indicate that for both adults and children there tends to be more pauses employed in the manuscript writing than in the cursive writing.

TABLE III.—COMPARISON OF TIME SPENT BY ADULTS IN PAUSES AT END OF STROKES

Subject	Time for manuscript		Time for cursive	
	22nds of a second	Per cent	22nds of a second	Per cent
1	128	25.4	79	21.1
2	175	33	79	19.4
3	286	50.2	164	39.7
4	202	38.3	178	35.7
5	260	49.5	244	43
6	368	56.2	209	38.4
7	463	59.1	233	43.6
8	190	36.6	138	31.9
9	118	28.7	93	24.1
10	219	34.8	157	30.9
Average	240.9	41.8	157.4	32.8

TABLE IV.—COMPARISON OF TIME SPENT BY CHILDREN IN PAUSES AT END OF STROKES

Manuscript			Cursive		
Subject	30ths of a second	Per cent	Subject	30ths of a second	Per cent
21	409	74.4	31	112	41.3
22	342	68.8	32	144	46.2
23	335	72.5	33	199	49.6
24	210	61	34	304	61
25	224	64.4	35	182	51.6
26	306	65.5	36	174	45.5
27	225	56	37	217	56.4
28	234	60.8	38	219	55.9
29	292	67.1	39	214	52.5
30	287	66	40	247	58.5
Average	286.4	65.7		201.2	51.9

It will also be noted that the difference in time spent in pauses between manuscript writing and cursive writing is not as great for adults as it is for children. This may be due to the fact that control increases with age, and consequently time is economized by reducing the pauses.

Table V gives a comparison between the number of pauses and the average length of time per pause for the complete records of each adult in manuscript writing and cursive writing. The table is to read as follows: For Subject 1 there were 112 pauses in the manuscript writing with an average length of 1.1 twenty-seconds of a second as compared with 71 pauses in the cursive writing with an average length of 1.1 twenty-seconds of a second. Very little difference is shown in the average duration per pause for the two types of writing. In the manuscript writing two of the writers, Subjects 6 and 7, spend on the average more than two twenty-seconds of a second per pause. In no other instance in either style of writing does any subject attain an average of this magnitude. For the group as a whole, the average is 1.6 units of time for manuscript writing as compared with 1.4 units for cursive writing.

A similar comparison is given for the children in Table VI. Subject 21, a manuscript writer, makes 93 pauses with an average length of 4.4 thirtieths of a second; Subject 31, a cursive writer, makes 58 pauses with an average length of 1.9 thirtieths of a second. The average

for the group is 3.5 units of time for manuscript writing as compared with 2.9 units for cursive writing.

TABLE V.—COMPARISON BETWEEN THE NUMBER OF PAUSES AND THE AVERAGE LENGTH OF TIME PER PAUSE IN TWENTY-SECONDS OF A SECOND IN WRITING OF ADULTS

Subject	Manuscript		Cursive	
	Number	Average	Number	Average
1	112	1.1	71	1.1
2	117	1.4	63	1.2
3	148	1.9	108	1.5
4	124	1.6	117	1.5
5	151	1.7	131	1.8
6	154	2.3	123	1.6
7	182	2.5	126	1.8
8	130	1.4	98	1.4
9	94	1.2	74	1.2
10	158	1.3	127	1.2
Total	1370	1.6	1038	1.4

TABLE VI.—COMPARISON BETWEEN THE NUMBER OF PAUSES AND THE AVERAGE LENGTH OF TIME PER PAUSE IN THIRTIETHS OF A SECOND IN WRITING OF CHILDREN

Manuscript			Cursive		
Subject	Number	Average	Subject	Number	Average
21	93	4.4	31	58	1.9
22	90	3.8	32	59	2.4
23	84	4	33	71	2.8
24	76	2.8	34	82	3.7
25	61	3.7	35	65	2.8
26	81	3.8	36	64	2.7
27	82	2.7	37	63	3.4
28	76	3.1	38	76	2.9
29	86	3.4	39	65	3.3
30	83	3.5	40	81	3
Total	812	3.5	..	684	2.9

Apparently then, the length of time consumed and the character of the writing are more largely dependent upon the number of pauses than upon the duration of pauses.

A question concerning manuscript writing which has been raised is whether time is lost in lifting the pen between letters. The issue has been investigated in the present study by counting from the speed curves the time spent in making strokes on the paper and in making strokes with pen raised from the paper. The length of stroke in each case was also recorded and the average length of stroke per unit of time calculated. This was done for three words, *pluck, and* and *zeal* in the case of adults and for one word, *school,* in the case of children. The results are presented in Tables VII and VIII. Table VII giving the data for adults is to be read as follows: Subject 1 spent 51 units of time in contact strokes with a total length of 100 units or an average of 1.98 fifths of a centimeter per unit of time as compared with 26 units of time in air strokes with a total length of 62.5 units or an average of 2.4 fifths of a centimeter per unit of time. It will be noted that for Subjects 4, 5 and 7 the average length of stroke per unit of time is slightly faster for the contact strokes. All other subjects show a higher average for the air strokes. The average for the group is 2.32 units of length for contact strokes as compared with 2.48 units for air strokes.

The results for the children are given in Table VIII which is to be read in the same manner as the preceding one. With the exception

TABLE VII.—COMPARISON OF SPEED OF STROKES MADE ON THE PAPER WITH THOSE PRODUCED IN THE AIR FOR MANUSCRIPT WRITING OF ADULTS

Subject	Contact strokes			Air strokes		
	Time in 22nds of a second	Length of stroke in ⅕ cm.	Average length of stroke per unit of time	Time in 22nds of a second	Length of stroke in ⅕ cm.	Average length of stroke per unit of time
1	51	100	1.98	26	62.5	2.40
2	57	107	1.87	25	59	2.36
3	41	99	2.41	21	57	2.71
4	50	132	2.64	16	35.5	2.21
5	44	102	2.31	22	46.5	2.11
6	43	107.5	2.50	19	51	2.68
7	43	103.5	2.40	24	57.5	2.39
8	50	130.5	2.61	20	58	2.90
9	40	98	2.45	21	52	2.47
10	59	134	2.27	33	85	2.57
Total	478	1113.5	2.32	227	564	2.48

TABLE VIII.—COMPARISON OF SPEED OF STROKES MADE ON THE PAPER WITH THOSE PRODUCED IN THE AIR FOR MANUSCRIPT WRITING OF CHILDREN

Subject	Contact strokes			Air strokes		
	Time in 30ths of a second	Length of stroke in ⅕ cm.	Average length stroke per unit of time	Time in 30ths of a second	Length of stroke in ⅕ cm.	Average length stroke per unit of time
21	23	39.5	1.71	9	21.5	2.38
22	28	54.5	1.94	17	42.5	2.50
23	23	46	2	11	25.5	2.31
24	27	52.5	1.94	13	31.5	2.42
25	21	50	2.38	10	28	2.80
26	31	65	2.09	15	39	2.60
27	31	67	2.16	16	42	2.62
28	24	49	2.04	23	39.5	1.71
29	27	64	2.37	16	40	2.50
30	27	50	1.85	13	30	2.30
Total	262	537.5	2.05	143	339.5	2.37

of Subject 28, the average length of stroke per unit of time is faster for air strokes than for contact strokes. For the group, the average is 2.05 units of length for contact strokes as compared with 2.37 units for air strokes.

These results indicate that there is apparently no loss of time in lifting the pen.

CONCLUSIONS

The results of the investigation show very clearly that there are differences in the movements of the two types of writing. These differences are found to be chiefly differences in the speed with which movements are made in writing letters of the two forms under consideration. No significant differences were found between manuscript writing and cursive writing in the manner of grasping the pen, in the sideward movement of the hand within words, in the method of moving the hand and arm along the line, or in the combination of hand movements, arm movements and finger movements executed while forming the letter strokes.

In neither type of writing is the movement made in forming words continuous. It is broken up into units with well-defined pauses between units. For example, in writing the word *pluck*, the pen may

trace the first stroke to the bottom of the *p* and then pause a short time before tracing the upward stroke, or it may trace a downward and an upward movement before a pause occurs. Sometimes a whole letter may be written without a pause but never a whole word. The portion of the movement made between pauses is a unit. The number of pauses in the two types of writing is somewhat different; the per cent of time spent in pauses at the end of strokes is greater in manuscript writing than it is in cursive writing. The reason for this greater length of time appears to lie in the form of the letters in the two types of writing. Pauses in writing come as a rule in the sharpest turns or changes in direction of the strokes. In manuscript writing some strokes are made on the paper and some are made in the air. most of the strokes made on the paper are produced in a downward direction and the strokes made in the air are produced in a forward and upward direction. At the end of strokes made in the air sharp changes in direction occur before strokes are made on the paper. The sharp changes in direction increase the number of pauses. The letters in cursive writing tend to be rounded in form and changes in the direction of strokes are not so abrupt. The number of pauses is thus reduced.

It has been shown in Tables I and II that cursive writing is faster than manuscript writing but the relative difference in speed between the two types of writing with age is not shown since the units of measurement are not the same in the two tables and the sentence written by the children was shorter than that written by the adults. Accordingly, Table IX has been constructed by reducing the data in Tables I and II to a common unit, number of letters per second. Table IX reads as follows: Adults write on the average 2.14 letters per second in manuscript writing as compared with 2.59 letters per

TABLE IX.—COMPARISON OF SPEED IN MANUSCRIPT WRITING AND CURSIVE WRITING

Group	Number of letters per second		
	Manuscript	Cursive	Difference
Adults	2.14	2.59	.45
Children	1.38	1.57	.19
Difference	.76	1.02	.26

second in cursive writing which indicates a difference of .45 letters per second in favor of cursive writing. The difference for the children is not so great, .19 letters per second in favor of cursive writing. There is an increase in speed with age for the two types of writing but the increase is less in the case of manuscript writing than it is in the case of cursive writing. The difference is .76 letters per second in manuscript writing and 1.02 letters per second in cursive writing or a difference of .26 letters per second in favor of the cursive writing.

Thus, adults write both manuscript writing and cursive faster than do children but there is a relatively greater increase in speed with age in the case of the cursive writing. Once more the influence of form is apparent. As an individual grows older his control increases with practice and consequently time is economized by reducing the number of pauses, but the form of the letters in manuscript writing prevents the reduction in the number of pauses to the extent possible in cursive writing.

In manuscript writing strokes made in the air are found to be slightly faster than strokes produced on the paper. Hence the relative slowness of manuscript writing is not due to lifting the pen. Length of pauses is found to be similar in the two types of writing but the number of pauses is greater in manuscript writing. The difference in speed is due partly to the greater number of pauses in manuscript writing and this feature is conditioned by the form of the letters.

Speed changes occur with greater uniformity in manuscript writing than they do in cursive writing. This characteristic does not appear to decrease or increase with age. Neither does it depend upon the individual, for several adults who maintained evenness of speed in manuscript writing failed to do so in cursive writing. The uniformity with which speed changes occur is apparently dependent to a large extent upon the form of the letters.

The results of the investigation indicate, then, that the differences between manuscript writing and cursive writing are chiefly in the speed changes which take place within the writing and are largely due to differences in the form of the letters.

SHOULD TEACHERS GIVE WARNING OF TESTS AND EXAMINATIONS?

GLENN R. PEASE

College of the Pacific

Probably every teacher has his own opinion respecting the value of the intensive review, or cramming, to which students almost invariably resort when given warning of a test. The real problem is a question of the value of such intensive review. In terms of time and effort, is such cramming worth while to the student? The problem must finally be settled in terms of immediate and delayed retention.

The stock argument against cramming for a test is aptly summed up in the well known adage, "Quickly learned, quickly forgotten." In other words, many argue that even though a student may be cramming pass a test with a high grade, he will retain little or nothing of the material so learned. There have been investigations which seem to uphold such a contention. Meumann[1] states that a rapid learner is also a rapid forgetter. Radosavljevich,[2] working in Meumann's laboratory, came to the conclusion that slow learners were the best retainers.

The majority of the evidence, however, seems to be on the other side of the question. Pyle[3] found a slight advantage on the side of the rapid learners. Busemann[4] states that in all probability a rapid learner is a good retainer. Gamble[5] found marked positive correlation between rapid learning and retention. Henderson[6] concludes that rapid learners are better retainers of the learned material than are slow learners.

Strayer and Norsworthy[7] say, "Quickly come, slowly go." Gates,[8] after a careful study, finds no support for Meumann's con-

[1] "Psychology of Learning." Baird translation, p. 169.

[2] "Das Behalten und Vergessen bei Kindern und Erwachsenen nach experimentellen Untersuchungen." Leipzig, 1907.

[3] Retention as Related to Repetition. *Journal of Educational Psychology*, 1911, Vol. XXII, p. 311.

[4] Lernen und Behalten. *Zeet fur angewandte Psychol.*, 1911, Vol. V, p. 211.

[5] "The Relation of Learning to Tenacity of Impression." Notes of an unpublished paper delivered at Washington Meeting of the American Association for the Advancement of Science, December, 1911.

[6] "Study in Memory." Psychological Review Monograph Supplement No. 23.

[7] "How to Teach." 1917, p. 81.

[8] Correlation of Immediate and Delayed Recall. *Journal of Educational Psychology*, 1918, Vol. IX, pp. 489–496.

clusion, but rather votes with Thorndike for a high positive correlation of abilities which would lead us to expect slow forgetting with rapid learning.

Lyon[1] concludes that subjects with high intelligence learn more quickly and retain better than subjects with low intelligence. Surely the preponderance of evidence for high correlation between quick learning and retention ought to overtop the weight of an adage, supported by a few early experimental findings on only a handful of subjects.

So far as the author has been able to ascertain, there has been little or no research respecting the value of intensive review, or as it is usually called, cramming. Such a study, to be worth anything, should be carried on under ordinary classroom conditions with considerable numbers of subjects rather than in a laboratory with a few subjects learning nonsense syllables. The study should give data for both immediate and delayed retention.

The subject first became of interest to the author several years ago while doing some graduate work in John Hopkins University, and later became the theme of his Master's research[2] at the State University of Iowa. Having had opportunity since that time to advance the study farther, it now seems worth while to review the earlier findings and to add the results of later research. The same technique has been followed throughout, so it had best be explained at once.

Test Procedure

The students making up the classes used in the study were given the Terman Group Test of Intelligence, Form A, with the time shortened one-fourth for college groups in order to scatter their scores more widely. On the basis of the intelligence test scores, the members of each class were divided into two paired groups, one of which became the experimental group, and the other the control group.

A one hundred item objective type test was constructed for each class, covering several months of the usual course work already completed by the students. On the day tests were to be given the author explained to each class that an experiment was being carried on in order to determine the value of cramming. The division of the class

[1] "Relation of Quickness of Learning to Retention." Archives of Psychology, No. 34, 1915.

[2] "The Effect of Cramming upon Retention, Immediate and Delayed." S. U. I. Master of Arts Thesis, 1927.

into experimental and control groups was explained, and they were told that one group was to be dismissed for cramming while the other group was to take the test at once. It was further made clear that the group dismissed was to spend at least an hour of intensive review in anticipation of a test at the next meeting of the class. They were warned that grades made in the tests would count on semester grades, but it was made clear that those who were in the control with no opportunity for review would not be penalized for lack of preparation, as grades were to be given on a comparative basis in each group. It was also explained that while the time spent in cramming would not count directly pro or con, an accurate record of the time was necessary for the purpose of the research, and that for this reason each member of the experimental group would report, at the time of taking the test, the number of minutes spent in review. A coin was tossed to determine which group was to serve as experimental group, and that group was dismissed while the control group took the test at once.

Students were allowed all the time necessary for the completion of the test up to the end of the period, fifty minutes. Most of them finished their tests in from twenty to thirty minutes. Test papers were collected immediately upon completion, so that there was no opportunity for copying of items. Passing on of information from control to experimental group was further guarded against by explaining that such procedure would invalidate the research, and by putting the students on their honor not to pass on any information respecting the nature of the test.

Six weeks later, with no class reviews nor further testing over the ground covered, the identical test was repeated as a surprise test for both groups in order to measure the amount of delayed retention.

Results

The first study used a total of 204 in each group, of whom 302 were college students, and 106 were high school pupils. The mean cramming time for the experimental group was 86.3 minutes. The mean lead of experimental over control group in the first test was 11.1 points in the hundred item test. In the re-test six weeks later, the experimental group still held a lead of 6.3 points over the control group.

The real significance of such differences is determined only when treated statistically. When the PE of difference is found by the formula

$$PE_{dif.} = \sqrt{(PE_m)^2 (PE_m)^2}$$

we are able to state the significance, for when the difference is as much as three times the PE of difference, it is considered statistically significant. Such treatment of the differences show that they are statistically significant both for the test and re-test in every class having as many as fifteen paired cases. Even in the classes where the difference was not significant, in no case does the experimental group fail to lead the control group in both the test and the re-test.

The conclusion drawn from the above findings is that, from the standpoint of the student, it pays to cram, not only for the purpose of raising the score in the immediate test, but for any test occurring up to six weeks later than the time of cramming. Eighty-six minutes of cramming raised the average grade a total of eleven points, which would make considerable difference in a letter grade for many pupils. Also, over half of such advantage is still retained at the end of six weeks time.

The more recent studies of the value of cramming incident to warning respecting tests have been carried on in psychology classes only, due to the fact that the author's teaching is done in that field.

There were ninety-four in one class studied. These students were divided, as explained above, into experimental and control groups on the basis of Terman Group Test scores. With a mean cramming time of 89.5 minutes the initial test shows a lead of experimental over control group of 8.2 points with a PE of difference of ± 2.76, which means a significant difference. The re-test given six weeks later shows a difference in favor of the experimental group of 4.8 ± 2.25, not a significant difference.

In another class there were 82. These students were paired as above explained, and given the tests in the same manner except that the retest was given twelve weeks later. On the initial test the experimental group, with a mean cramming time of 85.12 minutes, shows a mean lead of 17.07 points over the control group. The PE of difference is ± 3.325 showing it to be a statistically significant difference. The retest twelve weeks later shows a mean lead of experimental over control group of only 2.7 points with a PE of difference of ± 3.174, which is not a significant difference.

Some interesting conclusions might be drawn from these comparisons. The group given the retest six weeks after the initial test shows a mean Terman Group Test score of 139.1 while the group given the retest twelve weeks later shows a mean Terman Group Test score of 163.4. The group with the inferior intelligence crammed a slightly

longer time but raised their scores an average of only eight points compared to the other group's seventeen points. From these data it would appear that the value of cramming is in direct proportion to the intelligence of the one doing the cramming.

A further conclusion drawn from the data is that while something like a quarter to a half of the advantage gained by cramming is retained for six weeks, only about one-eighth of it remains at the end of twelve weeks. Whether these figures would have been the same had the two classes had the same level of intelligence is an unanswered question.

It seems to the author that some worthwhile research could be carried out along this idea in various fields of secondary and college training. If cramming makes for real learning, a goodly portion of which is retained for considerable lengths of time, the teachers should give warning of tests. So far as the field of psychology is concerned, the author considers that the proof for the value of cramming is clear, therefore the consequent value of giving warning respecting tests and examinations.

THE RESTANDARDIZATION OF IQ'S OF DIFFERENT TESTS

MINNIE L. STECKEL

University of Chicago

It is impossible, at present, to obtain a single group intelligence test which is practicable for all grades from I to XII inclusive. In making a survey of the intelligence of children of an entire school system, it is necessary, therefore, to select a number of group tests from those on the market in order adequately to test the various age groups.

According to the standardization claimed for such group tests, the IQ's of the various tests should be directly comparable; that is, a child making an IQ of 90 in one test should have the same ability as children making an IQ of 90 in any other test and the ability represented by an IQ in one test should be equal to the ability represented by the same IQ in any other test.

It is the purpose of this paper to show; (1) that the IQ's of several well known tests are not directly comparable and (2) to present a method by which they may be equated so that a direct comparison is possible.

The data for this study were obtained from 10,779 children in the public schools of Sioux City, Iowa, during the school year 1927–1928. They include the records of children from grades I to XII inclusive. The tests used for the various grades are as follows:

Tests	Grades	Number of Children
Kuhlmann-Anderson	Jr. I to Jr. III	2587
National	Sr. III to Sr. IV	1757
Otis intermediate	Jr. V to Sr. VIII	4301
Otis advanced	Jr. IX to Sr. XII	2134

Since (1) several thousand cases are used for each test; (2) the range of intelligence is essentially that for which each of the tests was constructed, (3) and the groups tested are drawn from the same population, we may consider the data adequate to establish a correspondence in general among the various tests in question.

Diagram I shows the distributions of IQ's of the four tests drawn upon the same base line. The test which was given to a child was

determined only by his age so that we should expect the distributions of intelligence quotients to be practically the same for the four tests or at least that the mean quotients should be the same. If there is any validity to the claim of constancy of IQ during the early school years then the four distributions in our diagram should be identical. Several thousand of the children in these distributions are siblings and we can think of no selective factor in this situation that would cause either the mean intelligence or its dispersion to vary from one age group to another if we assume the approximate constancy of the

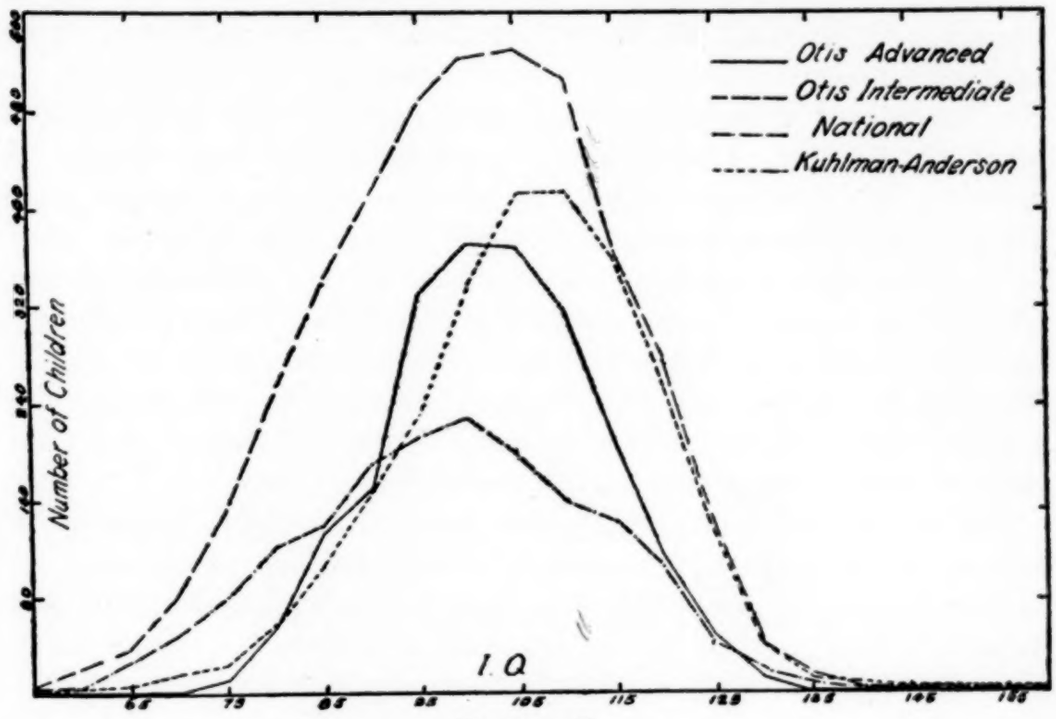


DIAGRAM I.

IQ during the school years. Therefore, a child who receives an IQ of 90 on the Kuhlmann-Anderson test at age seven should at ages nine, twelve and sixteen, if he took the National, the Otis Intermediate, and the Otis Advanced tests, respectively also receive an IQ of 90 on each of these tests.

But the four distributions are markedly different. This is probably easily explained by the facts that the four tests were standardized on different populations in different localities. The mean intelligence of eight year old children in one community is assigned an IQ of 100 in one of the tests, and the mean intelligence of eight year old children in another community is assigned an IQ of 100 on another test. Now

it is clear that if the two populations in different localities are not of equal mean intelligence, then the tests are not directly comparable as to their IQ's. We are assuming that the tests measure approximately the same type of intellect.

Whenever several tests are standardized on different populations which are not known to be equal in mean intelligence and in dispersion this discrepancy is likely to be found. When two such tests are given to the same child they will give different IQ's to the same child even though the two tests measure the same type of intellect. Other things being equal, preference should be given to the test which has been standardized on the largest population. This preference should hold even though among the tests compared are individual tests. In the latter case examiners usually give preference to the individual tests which often is incorrect. If one individual test such as the Stanford Binet and one group test have been given to a child and if the two intelligence quotients disagree, preference should be given to the test which has been standardized on the larger population. There is no reason why the particular population which happened to be used for standardizing the Stanford Binet should be set up as a standard for judging all other tests. It may very well happen that a group test is more valid than some individual test if the group test has been standardized on the larger population.

In the present case we have four group tests which disagree. All four of the tests have been given to school children in the same city so that we can be reasonably sure that the four populations are comparable. This enables us to restandardize these four tests on the same basis so that they should give comparable results if any two or three of them should be given to the same child, or in case the intelligence quotients of children taking different tests are to be compared.

The best criterion of brightness of a child is his relative standing among children of his own age. To make a comparison among different age groups we must first reduce the scores to comparable units. We are able to make both of these comparisons by means of percentile ranks. The assumption underlying a comparison made in this way is that the tests compared measure essentially the same thing. Diagram II shows the percentile ranks for the four tests. If the IQ's were comparable without equating, the four curves would coincide exactly. Since, however, this is not the case we may read directly from the chart the IQ in each of the tests which corresponds to a certain percentile rank, and thus equate them. For convenience

in making this transmutation, Table II has been constructed from the diagram. By use of this table we can readily read the IQ in one test and the corresponding IQ's in the other tests. For example, a child receiving an IQ of 90 in the Kuhlmann-Anderson test would, if he took the National test receive an IQ of 78, and a child receiving an

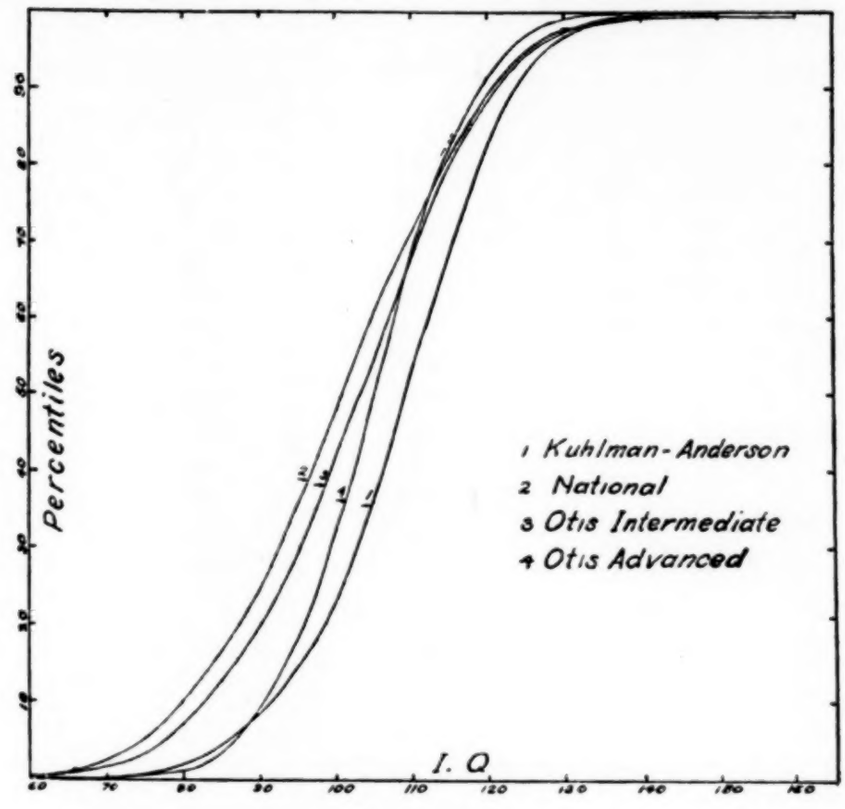


DIAGRAM II.

IQ of 90 in the National test would have the same ability as a child receiving an IQ of 98 in Otis Advanced test. Also, if we wish to make a comparison of groups of children which include the records from the various tests indiscriminately, we may, by use of this table, translate all the IQ's into the corresponding IQ's of one of the tests and thus make the comparison valid. When our comparison deals with abilities of children taking the same test the difference between IQ's may be considered a real difference. As soon, however, as we compare abilities of children taking two or more different tests we must first restandardize the ratings of the tests in question before the differences found may be considered valid.

Table II of this study, because of the large number of cases and the similarity of the populations, may be used in general as a means

to equate the IQ's of the four tests included. It is desirable that other tests commonly used in school surveys be restandardized in a similar manner making possible a direct comparison of a relative large number of tests.

TABLE I.—DISTRIBUTIONS OF IQ'S OF THE FOUR TESTS

Class interval	Kuhlmann-Anderson	National	Otis intermediate	Otis advanced
55	1	4	7	0
60	4	6	21	0
65	3	32	37	2
70	16	49	80	1
75	24	80	153	11
80	58	120	248	54
85	107	139	338	132
90	163	194	431	168
95	230	213	507	329
100	339	227	545	370
105	411	200	549	367
110	412	160	525	312
115	356	145	364	208
120	252	104	277	115
125	132	44	147	50
130	53	23	45	13
135	17	10	19	2
140	8	3	4	0
145	1	1	2	0
150	0	2	1	0
155	0	0	1	0
160	0	1	0	0
Total	2587	1757	4301	2134

TABLE II.—PERCENTILE RANKS OF THE FOUR TESTS

Percentile rank	Kuhlmann-Anderson test	National test	Otis intermediate	Otis advanced	Percentile rank	Kuhlmann-Anderson test	National test	Otis intermediate	Otis advanced
1	77	67	67	81	51	109	101	104	105
2	80	70	71	83	52	110	102	104	106
3	83	72	74	85	53	110	102	104	106
4	85	74	76	86	54	110	102	105	106
5	86	75	78	87	55	111	103	105	107
6	88	76	79	87	56	111	103	106	107
7	89	77	80	88	57	111	104	106	107
8	90	78	81	89	58	112	104	106	107
9	91	79	82	90	59	112	104	107	108
10	91	80	83	91	60	112	105	107	108
11	92	81	83	91	61	112	105	107	108
12	93	82	84	92	62	113	106	108	109
13	94	83	85	92	63	113	106	108	109
14	95	84	86	93	64	113	107	109	109
15	95	84	87	94	65	114	107	109	109
16	96	85	87	94	66	114	107	109	110
17	96	85	88	95	67	114	108	110	110
18	97	86	88	95	68	115	108	110	110
19	98	87	89	96	69	115	109	111	111
20	98	87	90	96	70	115	109	111	111
21	99	88	90	96	71	116	110	111	111
22	99	89	91	97	72	116	110	112	112
23	100	89	91	97	73	116	111	112	112
24	100	90	92	97	74	117	111	112	112
25	101	90	92	98	75	117	112	113	113
26	101	91	93	98	76	117	112	113	113
27	102	91	93	98	77	118	113	114	113
28	102	92	94	98	78	118	113	114	114
29	102	92	94	99	79	119	114	115	114
30	103	93	95	99	80	119	114	115	114
31	103	93	95	99	81	119	115	115	115
32	103	94	96	100	82	120	116	116	115
33	104	94	96	100	83	120	116	116	116
34	104	94	97	100	84	121	117	117	116
35	105	95	97	101	85	121	117	118	117
36	105	95	98	101	86	122	118	118	117
37	105	96	98	101	87	122	118	119	118
38	105	96	98	102	88	123	119	120	118
39	106	97	99	102	89	123	120	120	119
40	106	97	99	102	90	124	121	121	119
41	106	97	99	103	91	124	121	122	120
42	107	98	100	103	92	125	122	123	121
43	107	98	100	103	93	126	123	124	122
44	107	99	101	103	94	127	124	125	122
45	108	99	101	104	95	128	125	126	123
46	108	99	102	104	96	129	126	127	124
47	108	100	102	104	97	130	128	128	125
48	109	100	102	104	98	132	130	130	128
49	109	100	103	105	99	135	135	135	135
50	109	101	103	105	100	150	165	160	140

THE INTERPRETATION OF CORRELATION COEFFICIENTS IN TERMS OF DEPARTURE FROM PERFECT CORRELATION

EDWARD A. LINCOLN

Harvard University

Correlation coefficients are sometimes interpreted on the basis of departure from perfect correlation of the individual cases in the two series of measures. This is done by finding the probable error of estimate from the formula $.6745\sigma\sqrt{1 - r^2}$ and the result tells the deviation exceeded by one-half of the cases. Thus if the correlation coefficient is .70, one-half of the cases will depart from perfect correlation by $.48\sigma$ or more.

It is the opinion of the writer that a more meaningful interpretation may be made on the basis of the per cent of cases which will vary from perfect correlation by certain more commonly used multiples of the standard deviation, as 0.5σ, 1.0σ, 1.5σ, etc. Otis has already prepared and published a table showing the displacement of individual cases in terms of the median deviation.[1] It seems preferable, however, to calculate the displacement in terms of the standard deviation.

This is done by starting with the probable error of estimate for the obtained coefficient, and from this calculating the desired constants. For example, suppose a coefficient of .50 has been obtained. $PE_{Est.} = .6745\sigma\sqrt{1 - r^2} = .5841\sigma$. That is, one-half the cases will depart from perfect correlation by $.5841\sigma$. How many cases will depart by one whole standard deviation or more?

$$PE_{Est} = .5841\sigma$$
$$1.71PE = 1.0\sigma$$

From Normal Curve tables, 1.7PE cuts off 37.42 per cent of the area

$$37.42 \times 2 = 74.84 = \text{area cut off on both sides}$$
$$100.00 - 74.84 = 25.16 = \text{per cent outside } \pm 1.0\sigma$$

The values in the accompanying table have been calculated in the same manner. In order to make the table more useful, there have been included other constants which are often employed for interpreting correlation coefficients, namely, the coefficient of alienation, k, the probable error of estimate, and the per cent of cases remaining within

[1] Otis, A. S.: "Statistical Method in Educational Measurement." Yonkers: World Book Co., 1925, p. 225.

the same median deviation of the two distributions according to the calculations made by Otis. With this table at hand, it is easy to interpret a coefficient in any one of the four ways for which the table gives the necessary data.

TABLE I.—CONSTANTS TO AID IN THE INTERPRETATION OF CORRELATION COEFFICIENTS

Value of r	Value of k	PE of estimate (σ)	Per cent in same median deviation	Per cent of cases which change by certain multiples of σ			
				$.5\sigma$	1.0σ	1.5σ	2.0σ
.00	1.00	.67	19	62	32	13	5
.05	1.00	.67		62	32	13	5
.10	1.00	.67	20	62	32	13	4
.15	.99	.67		61	31	13	4
.20	.98	.66	21	61	31	13	4
.25	.97	.65		60	30	12	4
.30	.95	.64	22	60	29	12	4
.35	.94	.63		60	29	11	3
.40	.92	.62	24	58	28	10	3
.45	.89	.60		58	26	9	3
.50	.87	.58	26	56	25	8	2
.55	.84	.56		55	23	7	2
.60	.80	.54	29	54	21	6	1
.65	.76	.51		51	19	5	0.9
.70	.71	.48	34	48	16	4	0.5
.75	.66	.45		45	13	2	0.3
.80	.60	.40	41	41	10	1	0.1
.85	.53	.36		34	6	0.4	
.90	.44	.29	55	25	2	0.1	
.91	.41	.28		23	2		
.92	.39	.26		20	1		
.93	.37	.25		17	0.7		
.94	.34	.23		14	0.3		
.95	.31	.21	71	11	0.1		
.96	.28	.19		7			
.97	.24	.16		4			
.98	.20	.13	91	1			
.99	.14	.10					
1.00	.00	.00	100				

RESULTS OF REPEATED MEASUREMENT OF PUPIL ACHIEVEMENT

GERTRUDE HILDRETH

The Lincoln School of Teachers College

Results of standardized achievement tests afford data for the study of pupil progress in learning when the tests are given to the same children repeatedly for a period of years, provided the test items have a wide range of difficulty. The practical value of such data for the improvement of school practice, particularly with reference to provision for individual differences, pupil classification, and educational guidance, is apparent.

At the Lincoln School of Teachers College, the Stanford Achievement tests have been given annually since 1923 to pupils in Grades II to VIII for the purpose of determining the educational status and progress of pupils in reading, arithmetic and spelling. Alternate forms of the test have been given each year about the middle of May. The result is a series of repeated measures of individual pupils which forms the basis for this study.

The Stanford Achievement test is not scaled in absolute units of measurement comparable to the units of a meter stick. The unit of measurement at any one point on the achievement scale is not necessarily equal to that at any other point. The result is therefore a relative measurement rather than a true measure of growth in achievement.

The mental ability of all pupils has been measured with the Stanford Binet test administered at the time of each pupil's admission to the school and at intervals of about three years thereafter.

Due to changes in pupil personnel and occasional absences at the testing period the number of cases available is only a fraction of the school population. The study has been limited to pupils for whom data for a period of five years or longer are available. There are forty-two pupils with complete records for a five year period in Grades II to VI, forty-seven pupils for whom there are records of initial and final scores for the same period, and seven pupils with complete records for a seven year period in Grades II to VIII.

The significant facts may be classified as follows:

1. Group tendencies in growth in achievement.
2. Individual growth curves.

3. The relation of initial to final scores in the series.
4. Correlation of subjects.
5. Relation of achievement to capacity.

Group Data

Figures 1, 2, and 3 indicate the distribution of scores, the median scores, age and grade equivalents, and subject quotients in reading, arithmetic and spelling for the group of forty-two pupils who had the

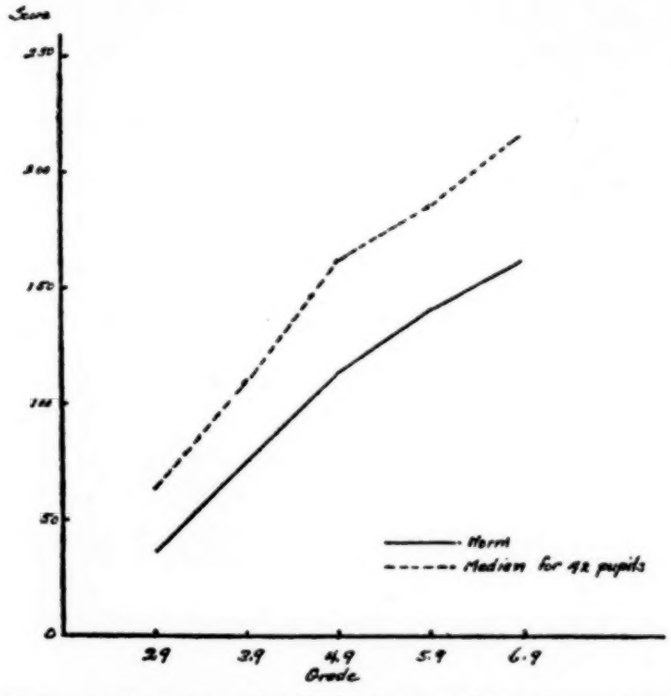


Fig. 1.—Median reading scores for forty-two pupils at successive grade levels.

Grade	2.9	3.9	4.9	5.9	6.9
Median score	64	110	163	187.7	216
Reading age	9-5	10-9	12-8	14-4	15-10
Grade equivalent	3.6	4.8	7.0	8.4	10.0
Reading quotient	119	121	128	131	133

test five years in succession in Grades II to VI. The graphs show the median curve for the group and the grade norms for the test.

These three curves, based on raw scores having different scale values are not directly comparable. They become comparable when the yearly increments are expressed in terms of per cents of maturity. These percents are given in Figure 4. The norms for the test converted into per cents of maturity differ slightly for the three subjects,

but the differences are too slight to require three different curves for norms.

The median curves in the different subjects show regularity of progress toward maturity from comparatively low initial scores. Rate of growth is similar in the different subjects, but status in reading is higher than that for arithmetic and spelling throughout the series.

The norms have been included to show the relationship between actual growth as measured by repeated tests of the same individuals,

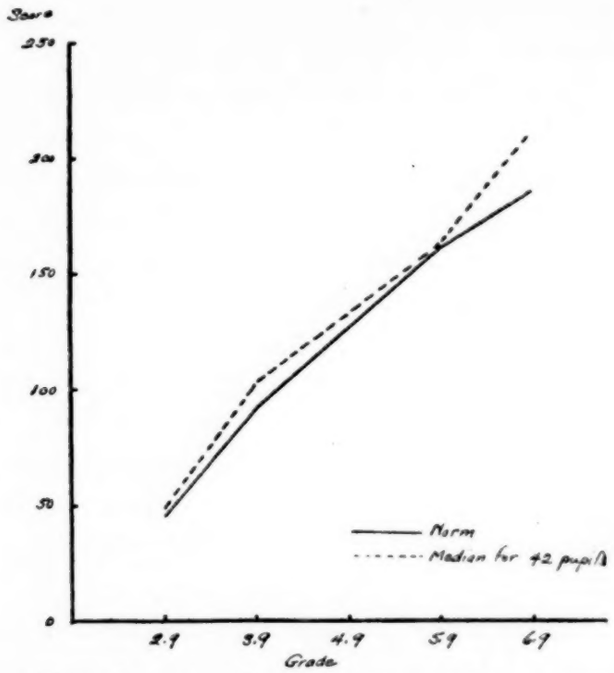


FIG. 2.—Median arithmetic scores for forty-two pupils at successive grade levels.

Grade	2.9	3.9	4.9	5.9	6.9
Median score	46.3	103.3	131.4	164	210
Arithmetic age	8-2	10-1	10-11	11–10	13-8
Grade equivalent	2.8	4.2	5.0	6.0	7.9
Arithmetic quotient	103	113	110	108	115

and the accomplishment of large groups of school pupils at several different grade levels. The graphs indicate the general similarity between rate of growth as indicated by norms and actual growth data.

The educational quotient eliminates the factor of age from grade to grade and indicates rate of educational progress. The quotients indicate that gain in momentum has been more rapid in reading than in arithmetic. In spelling the quotient remains practically constant throughout the series.

No effort has been made to determine the number, nature, or potency of various factors producing the present growth curves. It can only be conjectured that direct teaching of skills, and environmental and mental maturity factors have been potent in producing the resulting curves.

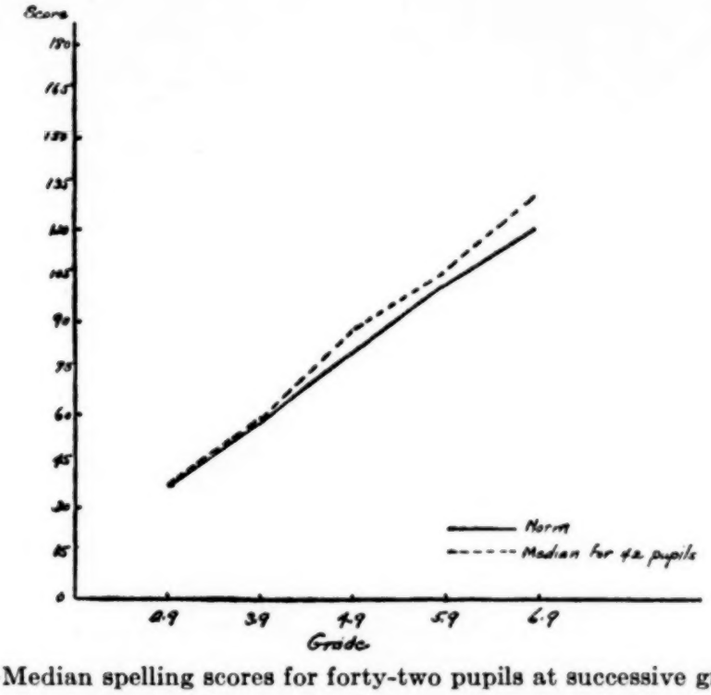


FIG. 3.—Median spelling scores for forty-two pupils at successive grade levels.

Grade	2.9	3.9	4.9	5.9	6.9
Median score	37.7	59	88.6	106.6	131.6
Spelling age	8-5	9-6	11-2	11-11	13-0
Grade equivalent	3.0	3.7	5.3	6.1	7.3
Spelling quotient	106	107	113	109	109

INDIVIDUAL GROWTH CURVES

Graphs showing the individual progress of seven pupils with series of test results for a seven year period are shown in Figures 5, 6 and 7. These curves are generally characterized by steady progress from initial to final status. Outstanding individual differences are observable. From the initial rank of these pupils final relative status could not very accurately have been predicted.

For further evidence concerning individual differences in rate of learning, three additional growth curves were constructed (Figure 8) from the data of Table I. These three pupils had respectively, the highest, lowest and median initial scores in reading of the group of forty-two pupils. It so happened that they were also the highest, lowest and median pupils in spelling and arithmetic. These three

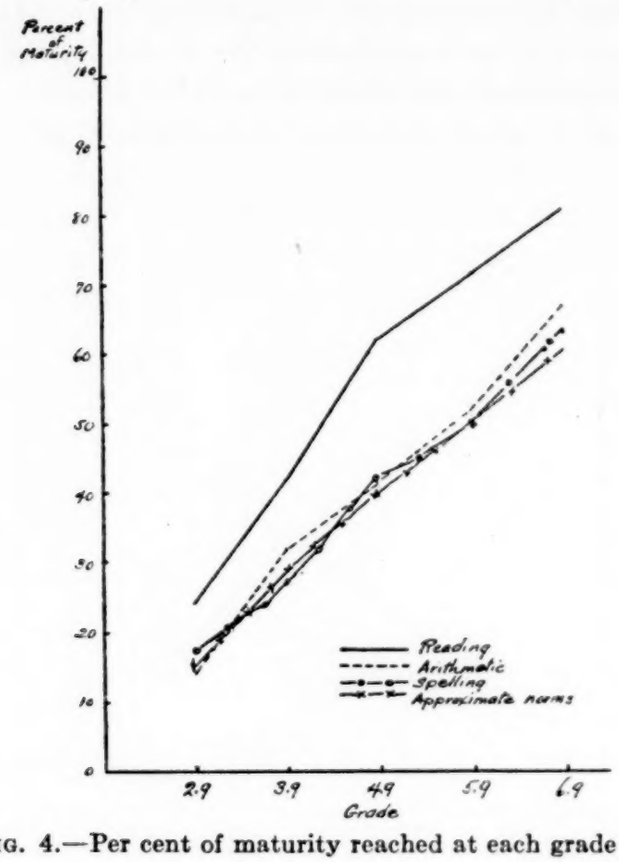


FIG. 4.—Per cent of maturity reached at each grade level.

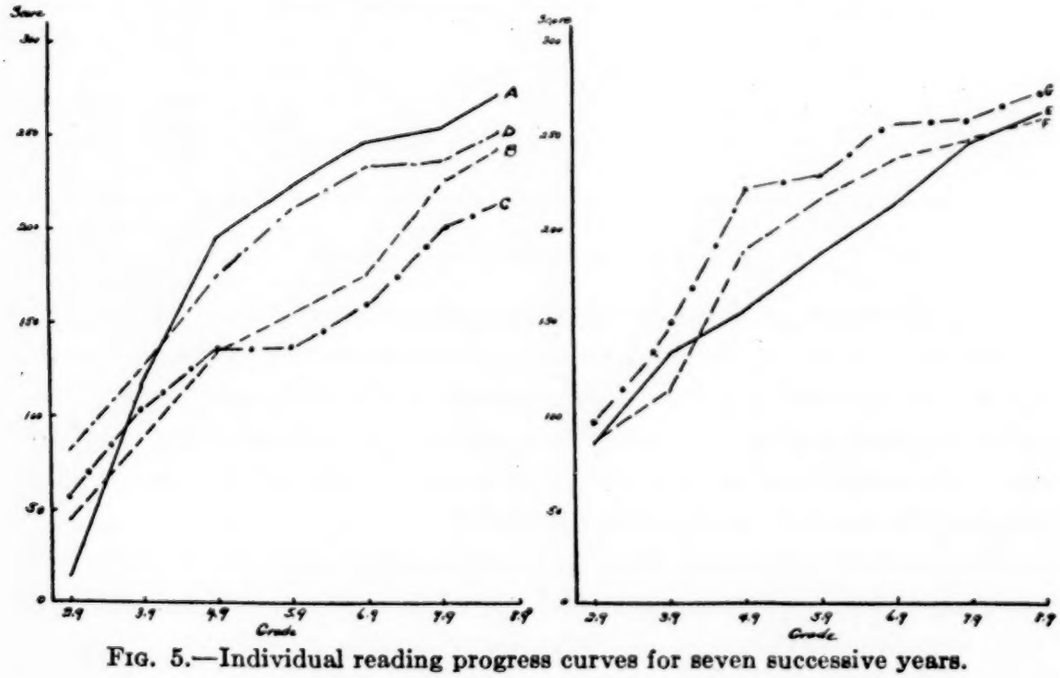


FIG. 5.—Individual reading progress curves for seven successive years.

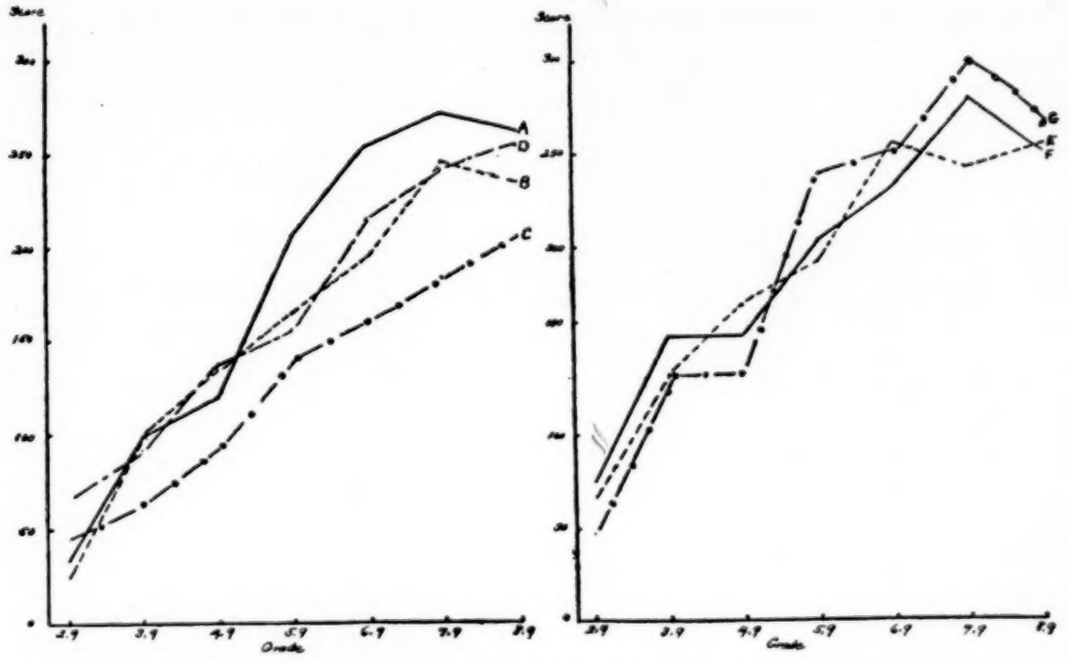


FIG. 6.—Individual arithmetic progress curves for seven successive years.

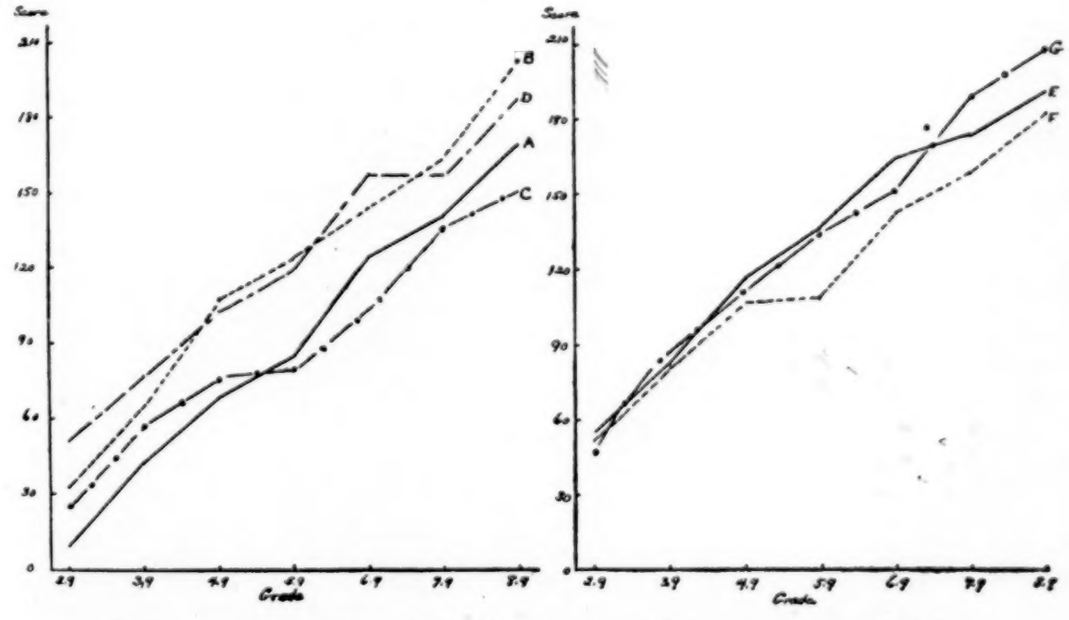


FIG. 7.—Individual spelling progress curves for seven successive years.

pupils therefore represent an average individual and the two extremes. The data indicate that these pupils retain their respective ranks at the

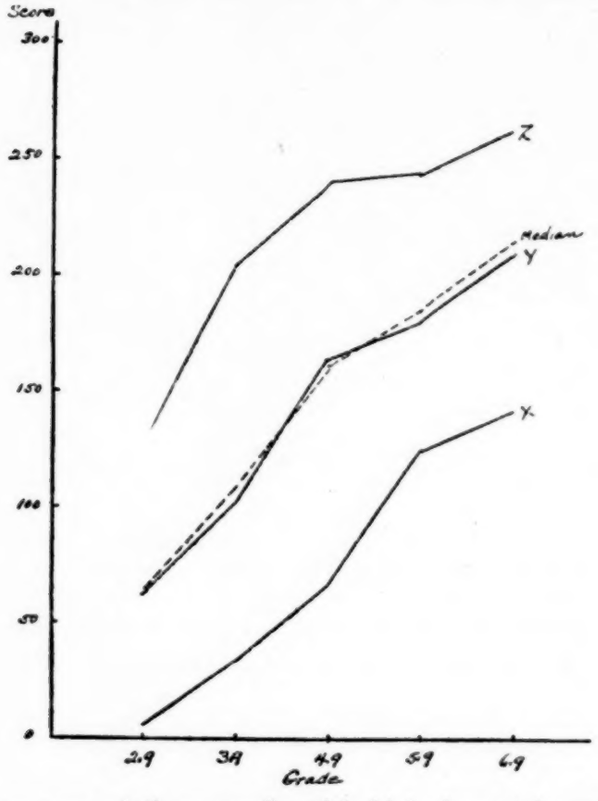


FIG. 8.—Reading progress of three pupils with high, low and median initial scores.

end of the five year period, and that rates of progress throughout the period are fairly even. The slow beginner is not the rapid learner at

TABLE I.—DATA FOR THREE PUPILS WITH HIGHEST, LOWEST AND MEDIAN INITIAL SCORES, RESPECTIVELY

Reading

Pupil	Grade				
	2.9	3.9	4.9	5.9	6.9
X	6	33	69	128	143
Y	63	114	166	182	211
Z	135	215	241	243	263
Median for forty-two pupils	64	110	163	187.7	216

the end nor *vice versa*. The curve of progress for the median initial scorer is very similar to the curve for the median of the entire group of

forty-two pupils. There is not in the case of these three pupils the occasional overlapping and fluctuation that marks the curves for the seven individual pupils with the longer series, pupils for whom initial status was less markedly differentiated.

Relation of Initial and Final Scores

The correlations for initial and final scores for forty-seven pupils were computed by the Pearson Product-moments method. The results follow:

Reading	.673 ± .05
Arithmetic	.478 ± .08
Spelling	.573 ± .07
Total score	.538 ± .07

The relation between initial and final score in the separate subjects is in all cases positive and appreciable, though not high enough for much accuracy in prediction.

The final status is a result of the initial status plus the gain the pupil makes in the interval. If a pupil has a high initial status it may be impossible for him to make a proportionate gain in score. His final status may therefore be relatively lowered by high initial status with resultant effect on the correlation coefficient.

Intercorrelation of Subjects

The relation of initial and the relation of final scores in reading, arithmetic and spelling is given in the following correlations:

Relation of Initial Scores

Reading and Arithmetic	.669 ± .05
Spelling and Arithmetic	.489 ± .07
Reading and Spelling	.729 ± .05

Relation of Final Scores

Reading and Arithmetic	.650 ± .06
Spelling and Arithmetic	.563 ± .07
Reading and Spelling	.753 ± .04

The intercorrelations between the three subjects are consistent when initial and final correlations are compared. Pupil success in one subject is positively correlated with success in another subject to

about the same extent near the beginning and near the end of the learning period.

Relation of Achievement to Ability

The average IQ of the group of forty-two children is 115. The AQ which expresses the relation of achievement to capacity, was computed for the median pupil in the separate subjects for each yearly interval. The results are as follows:

Reading

Grade	2.9	3.9	4.9	5.9	6.9
AQ	103	105	111	114	116

Arithmetic

Grade	2.9	3.9	4.9	5.9	6.9
AQ	90	98	96	94	100

Spelling

Grade	2.9	3.9	4.9	5.9	6.9
AQ	92	93	98	95	95

The AQ's increase from year to year in reading, but remain fairly constant in arithmetic and spelling. There is noticeable superiority of the accomplishment quotients in reading over accomplishment quotients in arithmetic and spelling.

The three pupils whose growth curves are shown in Figure 8, although chosen without reference to IQ were found to have widely differing ability. The IQ of pupil *X* is 97; of pupil *Y*, 118; of pupil *Z*, 150. *Z*'s IQ is the highest IQ found in the group of forty-two pupils and the IQ of the pupil *X* is the lowest. Pupil *Z* is also the youngest of the group and pupil *X* among the oldest. Pupil *Z* has skipped a year's work and pupil *X* has repeated a grade. Pupil *Y* has progressed regularly at the rate of a grade a year.

The relation of ability to achievement was studied further in the larger group of pupils through selecting groups of pupils differing significantly in IQ. Achievement data were tabulated for all pupils with IQ's over 120 (to be referred to as Group I) and for all pupils with IQ's below 110 (Group II). In Group I the median IQ for the seventeen cases was 126. For Group II with nineteen cases the median IQ was 106.

The initial and final scores for these two groups are as follows:

	Grade 2.9	Grade 6.9
Reading		
Group I	38	197.2
Group II	93	235.5
Spelling		
Group I	34	124
Group II	51	146
Arithmetic		
Group I	39	192.5
Group II	65.5	238

The median chronological age of the brighter group at the first test is seven years six months; of the less capable group, eight years two months. When the factor of difference in chronological age is considered, the differences in achievement between the two groups are even more marked. These differences in achievement have fairly high significance but the small number of cases reduces the chances of certainty.

It may be suggested that the bright group excel the less capable group because their brightness has been recognized early and they have been stimulated to achievement commensurate with their ability. The facts of the case, however, are to the contrary. A disproportionate amount of effort to improve achievement has, as is the universal custom, been expended on the less capable pupils. They have almost without exception had extra individual drill and "remedial" work. They are undoubtedly better off than they would have been without such extra help, but the added pressure has not raised the scores of the slower learners to a point of equality with those of rapid learners.

Summary and Conclusions

1. This study reports achievement growth data for a group of pupils in Grades II to VIII who were measured annually with the Stanford Achievement test.

2. Medians for group growth curves indicate more rapid acceleration in reading than in arithmetic and spelling; and in the cases of reading, but not of spelling and arithmetic, marked superiority to the norms for the test. The acceleration from year to year is fairly regular. The EQ indicates that rate of growth is most rapid in reading.

3. Individual growth curves for seven pupils who had a series of seven annual tests indicate significant differences in rate of growth,

but little differences in the nature of the growth curve. These seven pupils do not maintain the same respective status throughout the growth period. When the curves for the three individuals who rank respectively highest, lowest and median on the initial reading test are compared they are found to maintain their respective ranks and their rate of growth throughout the five year series. This is found to be equally true for their growth in spelling and arithmetic.

4. Correlations for initial and final scores in the separate subjects are positive and appreciable, but not high enough for great accuracy in predicting final from initial score.

5. The intercorrelations of subjects are also positive, and substantial, but not high. Correlations for initial status appear to be higher between reading and arithmetic, and reading and spelling than between spelling and arithmetic. The correlations are of similar size when intercorrelations for final status are computed.

6. The relation of capacity of the groups to achievement is indicated by the AQ's. These are higher in reading than in arithmetic or spelling. The IQ's of the three pupils of high, low and average achievement were respectively the highest, lowest and approximate median of the group of forty-two pupils from which they were chosen. In two groups of pupils selected on the basis of IQ so that the medians of the groups differ significantly, the brighter group is superior throughout the series in reading, arithmetic and spelling achievement.

AN EXPERIMENT ON THE RELATIVE EFFICIENCY OF STUDYING FRENCH VOCABULARY IN ASSOCIATED PAIRS VERSUS STUDYING FRENCH VOCABULARY IN CONTEXT

LOUISE C. SEIBERT

Goucher College, Baltimore, Maryland

The purpose of this experiment was to ascertain the relative efficiency of studying French vocabulary in English-French associated pairs versus studying the same vocabulary in some kind of context.

A series of twelve English-French words were learned by four different methods. Series A was learned in a list of twelve English-French associated pairs. Series B was learned in a series of twelve sentences, each one containing one of the words to be learned. Series C was learned half in context and half in associated pairs, and Series D was learned half in associated pairs and half in the context. These four different methods will be called Method P, Method C, Method CP, and Method PC, respectively.

To illustrate the kind of material used and the way it was used, we will suppose that the word "bit" is to be learned. In Series A, the English word will be given with its French associate, "bit" . . . "mors" and studied in this way. In Series B, the word will be used in a French sentence, chosen in such a way that the meaning of the word will be made clear by the sentence, thus: "On met *le mors* (bit) dans la bouche du cheval."[1]

Notice that the words in the sentence are all simple, and the meaning easy to get, both of the whole sentence and of the unknown word. The word to be learned was furthermore heavily underlined and its English translation was given immediately after. In Method C, the words were all studied exclusively in the sentences. In Method CP, the word was studied half of the time in the sentence, and half of the time in the associated list, while in Method PC, the word was studied first half of the time in the associated list, and the rest of the time in the sentence.

The purpose of these different methods was to follow, as closely as possible, the four methods employed most commonly by teachers and in the textbooks in presenting the vocabulary. Method P (associated pairs) is the straightforward method which has fallen very much into

[1] The bit is placed in the horse's mouth.

disrepute among progressive teachers. Method C, is the method of acquiring vocabulary entirely through the reading of the text, much as is acquired by the students through their casual reading. The translation of the word in English would represent the searching through the dictionary for the meaning of the unknown word. Method PC and CP, represent the familiar methods used by most of the French grammars. One will have a list of English-French words at the beginning of the lesson and will use these words in a context, the other will on the contrary, first present the words in the context, and study the words afterwards.

It will be well to bear in mind that our objective in this experiment is simply to find the relative value of these different methods as far as vocabulary learning is concerned, and that we do not consider whether other benefits, real or imaginary, may be derived from one method rather than from another.

One of the objections made to the learning of vocabulary by associated pairs, is that the words learned in this way are useless, inasmuch as the student is often unable to use them in context when he needs them. It was then important to ascertain whether this was true or not. By the same reasoning, it was also important to find out whether the words learned in context could be remembered independently of this context. This experiment deals exclusively with the knowledge of active vocabulary.

In order to find out whether words studied in context could be remembered separately from the context, and whether words studied in pairs could be used in a context, there were two recalls. In one recall, the student was given the list of English words, and their French associates were asked for. In the second recall (following immediately after the first one) the French sentence would be written at the board, with the word studied missing and replaced by its English translation. The student was asked to replace the French word in the sentence.

Through a slight error in procedure, the recalls were always given in the same way, *i.e.*, the recalls by pairs always preceded the recalls in sentences, so that the recall in the context could benefit from the first recall; furthermore, the sentences used in the original learning were the same as those used in the recalls. We did not want the recall by context to influence the recall by pairs, but by doing so, the procedure benefited the recall by context. A rotation in the procedure could have taken care of this but was not thought of at the time.

However the results were such, that in spite of these advantages, the recall most benefited was still the poorest in its results, and we could draw our conclusions with safety.

Very little work has been done on this particular subject. W. J. Grinstead tested the associational influence of context in fixing the meaning of new words encountered. He read the words in a German context and when an unknown word would present itself, he would look for its meaning in the dictionary. On the other hand, words given in a formal list were also looked up in the dictionary. He found that in the recall the words encountered in the context were remembered better.[1] This experiment, however, cannot be taken as conclusive, since it was only with one individual and no control of any sort was resorted to.

A group of sixty college students in second year French participated in this experiment. The group was divided into four sections of about the same number of students.

In order to counterbalance the possible effects of practice, the methods and lists were rotated in the following way:

Group I		Group II		Group III		Group IV	
List	Method	List	Method	List	Method	List	Method
A	P	B	PC	C	CP	D	C
B	C	A	CP	D	PC	C	P
C	PC	D	P	A	C	B	CP
D	CP	C	C	B	P	A	PC

Printed lists were distributed to the students. They were requested to study them aloud for nine minutes. The sentences were to be read in their entirety every time they wished to repeat the word to be learned. They were warned in advance that there would be a recall, and in what manner that recall would be given. In the case of the mixed methods (half sentence and half pairs) the students were given the two texts, and at the end of four minutes and a half changed quickly from the one to another. As the students were prepared for such a change, it took scarcely more than one or two seconds to do

[1] Grinstead, Wren J.: An Experiment in the Learning of Foreign Words. *Journal of Educational Psychology*, 1915, Vol. VI, pp. 242–245.

it. At the end of nine minutes, the texts were collected and the class work resumed.

The first recall was made at the end of fifty minutes. The English words or the French sentences were written on the board in a different order from that of the printed lists. The students had two slips of paper. On the first they wrote the French associate of the English word, on the other they wrote the missing French word in the sentence. They were warned not to make any alteration on the first slip, should they happen to recall something more in the following recall.

There were to be no relearnings, and it was thought that a certain amount of overlearning would be desirable to carry the material over for a period of forty-days. Taking the time of a previous experiment as a basis, we chose nine minutes for the initial learning.

The recalls were made at respective intervals of fifty minutes, two days, ten days and forty days.

Tabulating the Results

For each experiment, we have two sets of recalls, one for the words asked for in the context, the other for the words in associated pairs. The first recall will be called C, and the second P. It is well to have in mind that the P recall always preceded the C recall.

Table I.—Amount Retained after Fifty Minutes

	Method P		Method PC		Method CP		Method C	
	Pairs	Context	Pairs	Context	Pairs	Context	Pairs	Context
Total.....	2009	2004	1860	1865	1904	1900	1690	1717
Average.....	33.5	33.4	31.00	31.1	31.7	31.6	28.2	28.6
SD.....	3.84	4.20	5.61	5.49	5.13	5.31	7.62	7.35
$\left(\frac{SD}{\sqrt{n}}\right)$.....	.49	.54	.73	.71	.66	.68	.98	.95

Table I shows that the method of learning vocabulary by associated pairs gives the best results of the four methods. The learning in the context gives the poorest results, while the mixed methods give intermediate results, very similar to each other. The superiority of the method of pairs associated is quite marked, the average being 33.5 against 28.2 for the method by context. It is also interesting to notice how similar are the results for the two successive recalls,

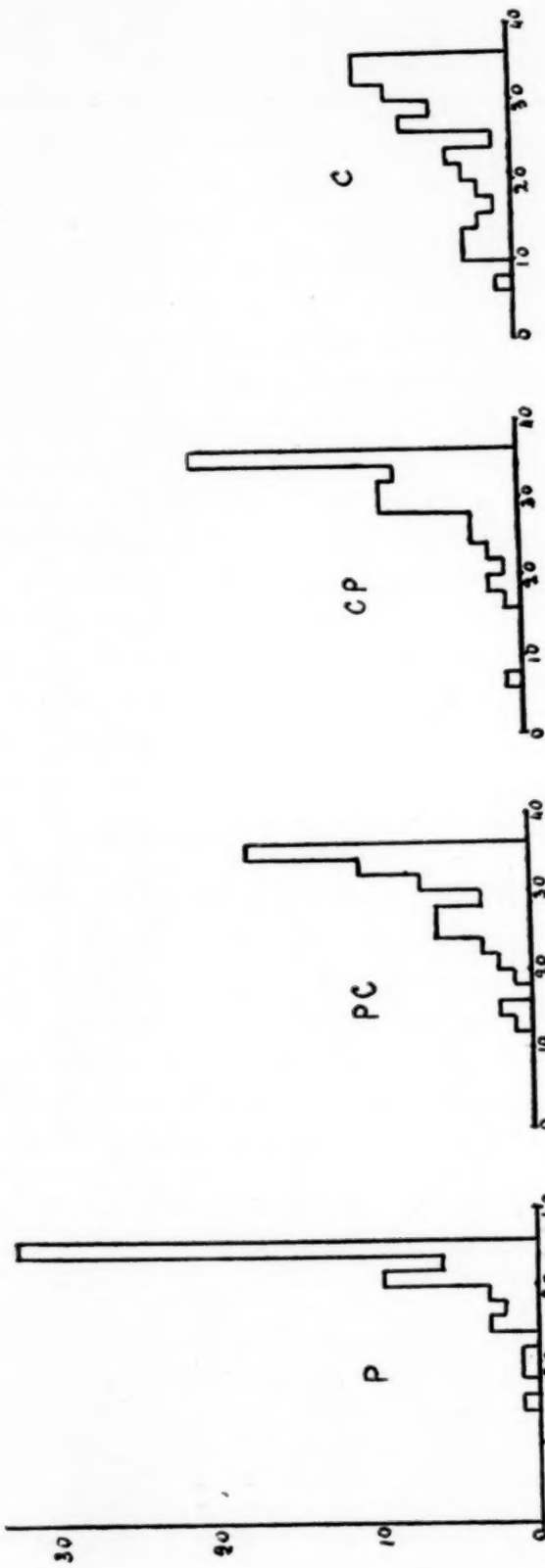


Fig 1 - Amount retained after 50 minutes (P recall)

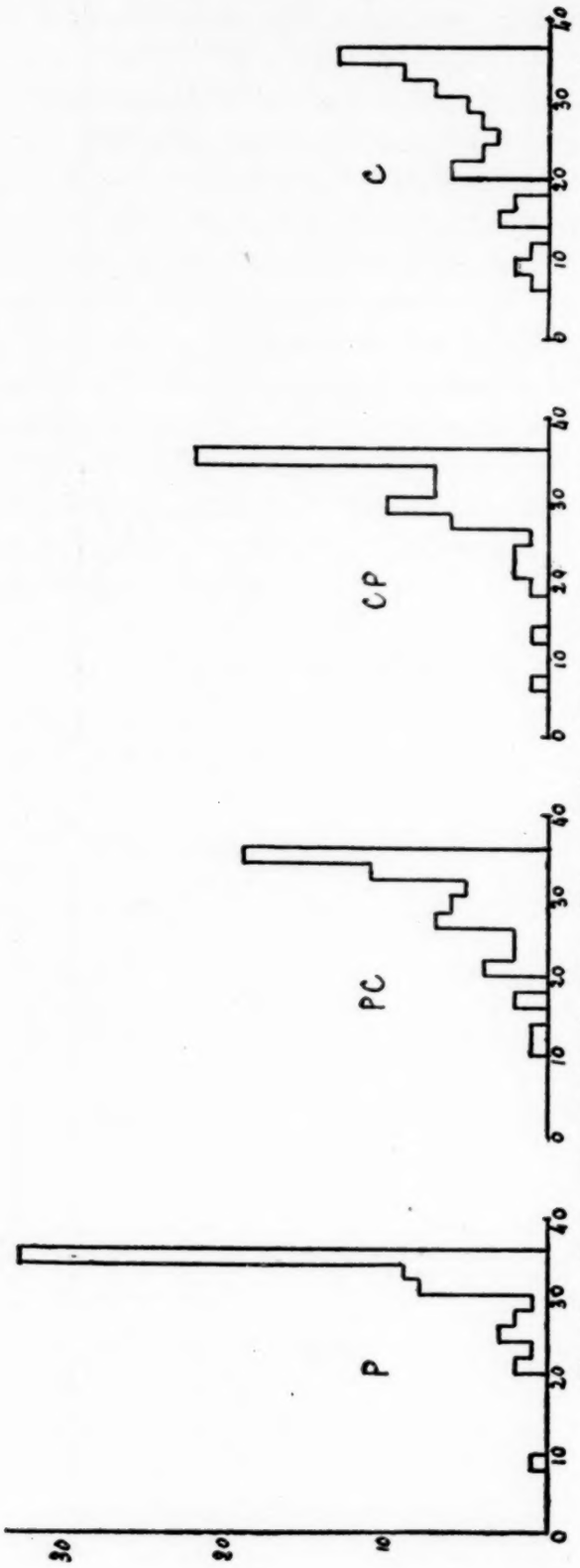


Fig. 2. Amount retained after 50 minutes (C recall)

although the recall by pairs was always given first. The only case where the results are somewhat dissimilar is in the method by context, which would tend to show that while the words learned by pairs are readily enough used in a context, the transfer from the context to the associated pairs is not so easy. However, the difference is very small in all cases, and we can say safely that whoever learned a word by the method of associated pairs could use it readily in a context and vice versa.

The SD of the distributions being very small, the averages are quite reliable as far as the number of students is concerned.

Figures 1 and 2 represent the accuracy distributions for the P and C recalls, respectively after an interval of fifty minutes.

The distributions, very compact for Method P, spread more and more as we pass to Methods PC, CP and C. The SD's for the P Method are respectively 3.8 and 4.2, while they are 5.61 and 5.49 for PC, 5.13 and 5.31 for CP and 7.62 and 7.35 for C.

The character of the distributions is also very interesting. The distributions for the recall of Method P shows a great number of perfect scores, in fact two-thirds of the total number of scores, while only one-third made a perfect score in Method C and one-half made a perfect score in the mixed methods. The SD of the P distributions are remarkably small.

The graphic representations for the two results are very much alike in character, although the recall in the sentence is slightly superior to the other. That it does not show any more superiority is rather surprising considering that the recall by context came after the recall by pairs, and that the same sentence was used all the time.

Table II gives the means and SD's of the accuracy distributions for all methods and both recalls, after an interval of two days.

TABLE II.—ACCURACY DISTRIBUTIONS AFTER AN INTERVAL OF TWO DAYS

	Method P		Method PC		Method CP		Method C	
	Pairs	Context	Pairs	Context	Pairs	Context	Pairs	Context
Total.....	1898	1904	1717	1729	1715	1721	1688	1714
Average.....	31.6	31.7	28.6	28.8	28.6	28.7	28.1	28.6
SD.....	4.95	5.58	6.15	6.42	5.91	6.90	6.96	6.51
$\left(\frac{SD}{\sqrt{n}}\right)$.....	.64	.72	.79	.82	.76	.89	.89	.84

The method of learning by associated pairs maintains its marked superiority after an interval of two days. The respective averages being 31.6 and 31.7 against 28.1 and 28.6 for the method of learning by context. We must note however that the loss for the C method is almost negligible while the average for the P method has dropped from 33.5 and 33.4 respectively to 31.6 and 31.7. The showing of the mixed methods is interesting. Their averages, which were superior to the average of the C method in the first recall, have now dropped to the level of the average of the C method.

Let us note also the very small superiority of the second recall (C recall) to the first recall (P recall) in Table II.

Figures 3 and 4 represent the distribution for the P and C recall respectively after an interval of two days.

The distributions, still very compact for Method P, widen from Method P to Method C. The distributions for the last three methods are strikingly similar. The P method shows the greatest number of perfect scores, especially the C recall. This difference of perfect scores in the two successive recalls shows that a second recall, while insufficient to add any more words, would probably tend to correct small mistakes in spelling or accentuation, thus changing the score from one to two points.

Table III gives the results for the recall after an interval of ten days, for both P and C recalls.

TABLE III.—ACCURACY DISTRIBUTIONS AFTER AN INTERVAL OF TEN DAYS

	Method P		Method PC		Method CP		Method C	
	Pairs	Context	Pairs	Context	Pairs	Context	Pairs	Context
Total	1915	1910	1693	1701	1713	1716	1647	1657
Average	31.9	31.8	28.2	28.3	28.5	28.6	27.4	27.6
SD	4.83	4.92	7.02	6.90	6.78	6.75	7.35	7.11
$\left(\frac{SD}{\sqrt{n}}\right)$	.62	.63	.90	.88	.87	.87	.94	.92

The method of associated pairs is still superior, the average for the recalls being 31.9 and 31.8 against 27.4 and 27.6 for the context method. The most striking fact about these results is that they are almost identical to the results obtained after an interval of two days. Method P shows a very slight increase. No relearning took place between the

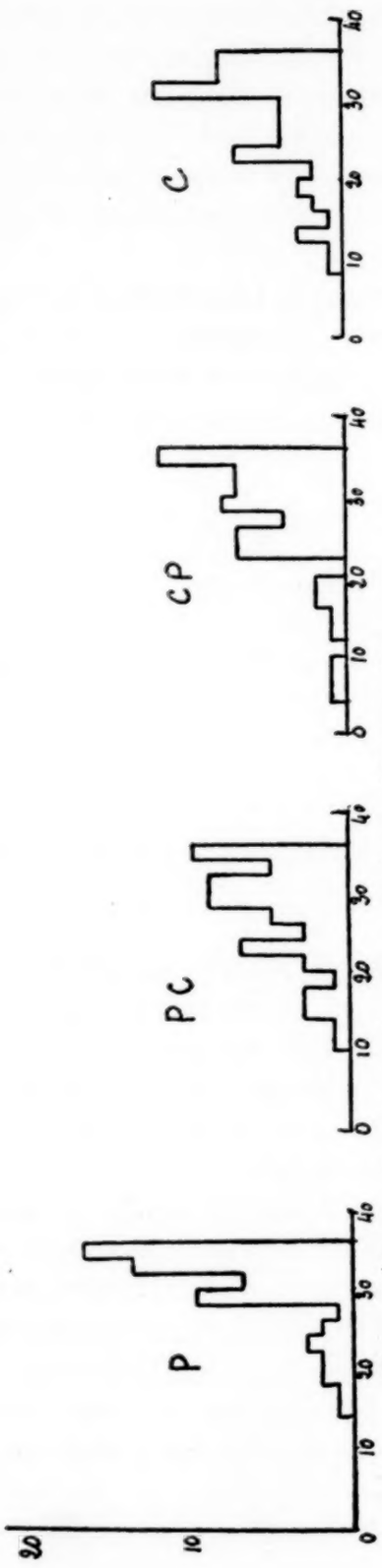


Fig. 3 - Amount retained after 2 days (P recall)

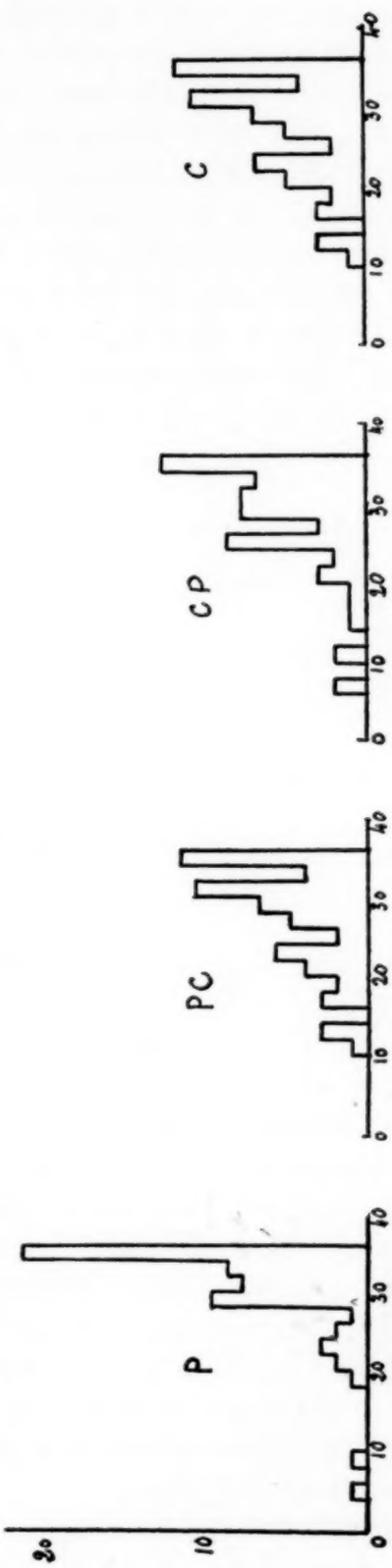


Fig. 4 - Amount retained after 2 days (C recall)

original learning and these recalls. On other experiments when a similar fact occurred, we attributed it to the fact that a relearning had taken place between the two recalls. But such does not seem to be the case. Method C shows the poorest results, and it is remarkable to see how the results of the mixed methods follow each other so closely.

Figures 5 and 6 represent graphically the distributions after an interval of ten days for P and C recalls.

The distribution for the P method remains remarkably compact, and still shows a great many perfect scores. The three other distributions are more widely spread and are very similar to each other.

Table IV gives the results for the recalls after an interval of forty days.

TABLE IV.—AMOUNT RETAINED AFTER AN INTERVAL OF FORTY DAYS

	Method P		Method PC		Method CP		Method C	
	Pairs	Context	Pairs	Context	Pairs	Context	Pairs	Context
Total.....	1493	1532	1316	1350	1324	1398	1247	1290
Average.....	24.9	25.5	21.9	22.5	22.0	23.3	20.8	21.5
SD.....	7.50	7.35	7.50	7.26	7.95	7.89	8.01	7.89
$\frac{(SD)}{\sqrt{n}}$.....	.97	.94	.98	.93	1.02	1.01	1.03	1.02

After an interval of forty days, the results remain consistent with those of earlier recalls. The method by associated pairs is still superior by a tangible amount to the method by context, the averages being 24.9 and 25.5 against 20.8 and 21.5 respectively. The mixed methods, however, have regained somewhat and stand in the middle ground just as they did after a fifty minute recall.

There is a constant improvement in the second recall (C recall) over the first recall (P recall). It is the first time that we observe any noticeable improvement. Whether this is due to the fact that a second recall will always be better than the first, or because the sentence has helped in the recall we have no way of ascertaining. Judging from the results of the other recalls, where the sentence did not help at all, we are inclined to attribute the difference to the fact that the C recall came in the second place.

Figures 7 and 8 show the graphical representations of the accuracy distributions after an interval of 40 days.

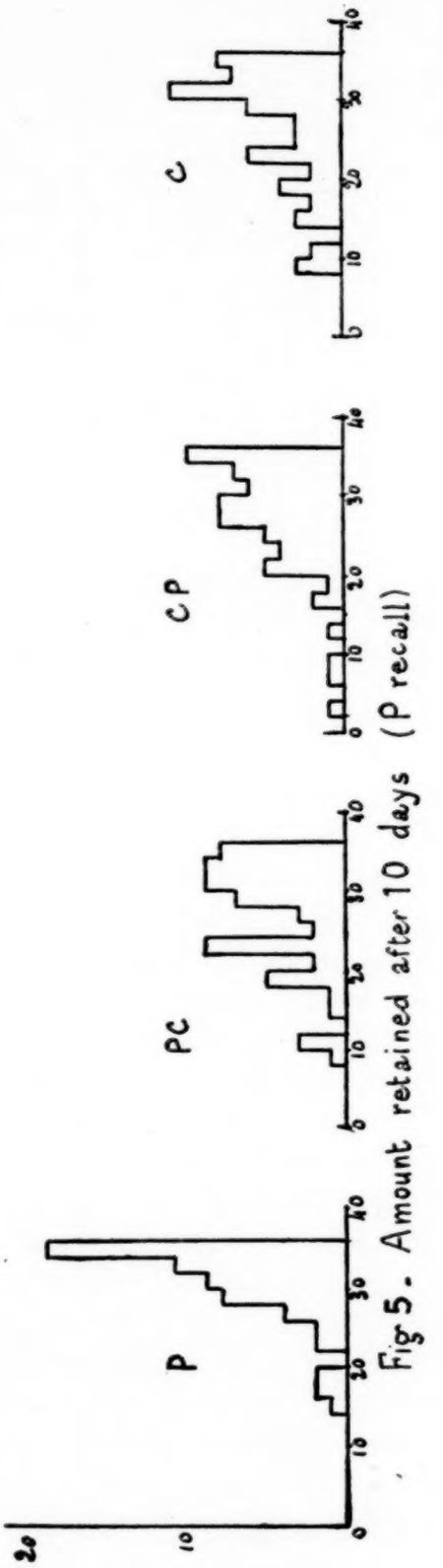


Fig 5 - Amount retained after 10 days (P recall)

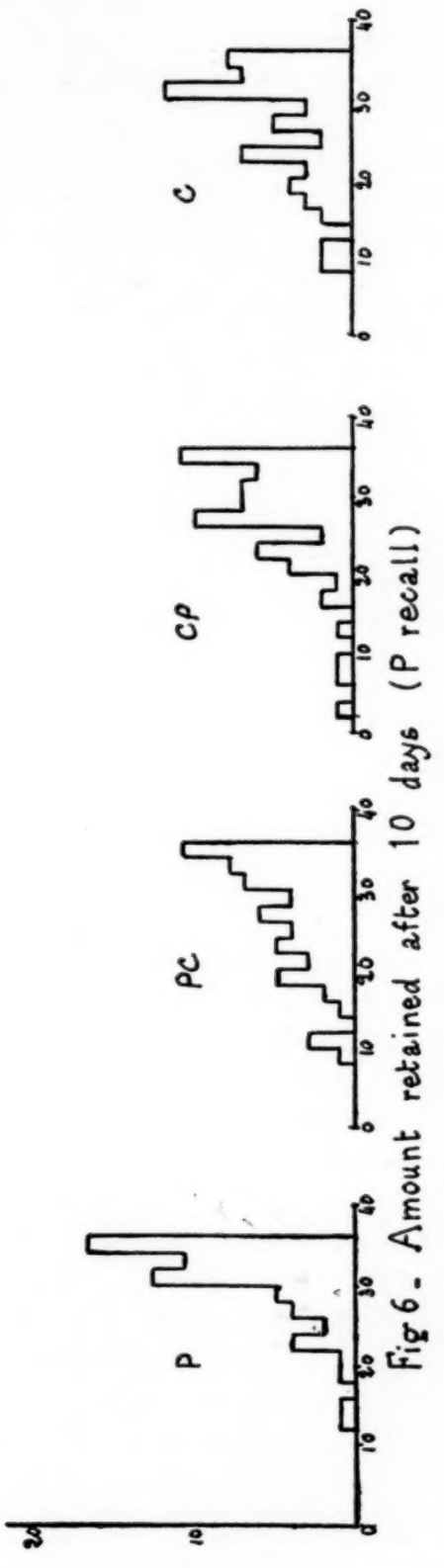


Fig 6 - Amount retained after 10 days (P recall)

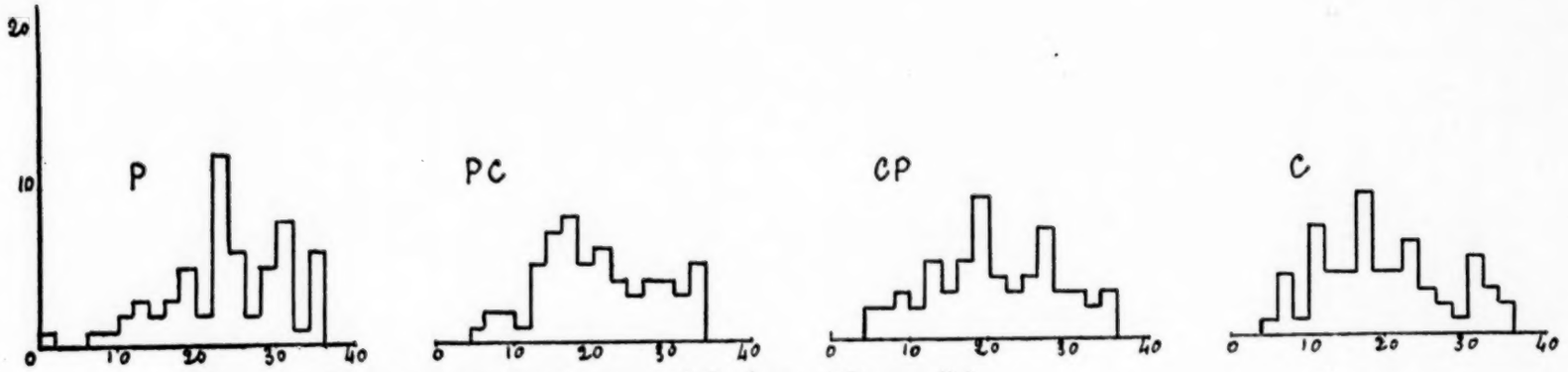


Fig. 7 - Amount retained after 40 days (P recall)

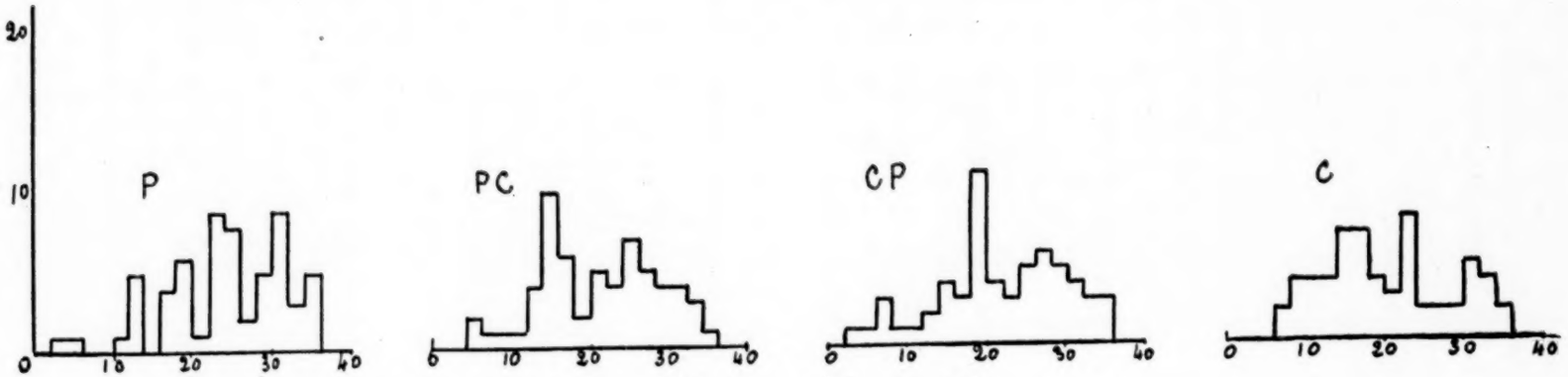


Fig 8. Amount retained after 40 days (C recall)

The first method has lost its compactness, and the four distributions look very much alike. The value of the SD is almost the same for all, and the curves have flattened considerably.

Per Cents

If we interpret the results into per cents of the perfect scores, the results will be clearer. The maximum score[1] is 2160 and the amount retained is given in Table V.

Table V.—Amount Retained in Per Cents of the Perfect Score

Recalls	Method P		Method PC		Method CP		Method C	
	Pairs	Context	Pairs	Context	Pairs	Context	Pairs	Context
First.....	93.01	92.78	86.12	86.35	88.15	87.97	78.25	79.50
Second.....	87.88	88.15	79.49	80.05	79.40	79.68	78.16	79.40
Third.....	88.66	88.43	78.38	78.75	79.40	79.41	76.25	76.72
Fourth.....	69.13	70.93	60.94	62.50	61.30	64.73	57.74	59.73

The results show that the amount retained after fifty minutes by the P method is 93.01 per cent and 92.78 per cent of the perfect score while it is 78.25 per cent and 79.50 per cent of the perfect score for the C method. In other words, the P method surpasses the C method from 14.76 per cent to 13.28 per cent of the perfect score.

Comparing the results after forty days, we find that the method of pairs associates gives results superior to the context method by 11.20 per cent of the perfect score. This superiority maintains itself through all the recalls. Another remarkable fact about these results is that all the recalls are so consistent with the original recalls and with each other. Roughly speaking, there is for every method a loss of about 20 to 25 per cent of the perfect score between the immediate recall and the final recall after forty days.

There is also a remarkable consistency between the two consecutive recalls (P recall) and (C recall). Nowhere is there seen a superiority marked enough to attribute it to other causes than to the very natural improvement that a second recall would show over the first. Warner Brown says: "A second recall contains more items than the first, because repetition tends to fix the items of the first recall, while chance tends to introduce new items into the second recall, in addition

[1] For the method of scoring see: An Experiment in the Learning of French Vocabulary by Louise C. Seibert. *Journal Educational Psychology*, May, 1927.

to those retained from the first recall."[1] We did expect however that the sentence would prove a bigger help in the recall of the words than the English associate, but there is not a trace of this influence.

The per cent retained at the end of two days and at the end of ten days are very similar. Evidently the amount of forgetting during this period is very small. We find it to be so in other experiments done on the same subject. Also the amount retained is very high; from ninety-three per cent to sixty-nine per cent after fifty minutes to an amount varying from seventy-nine to fifty-nine per cent of the perfect score after an interval of forty days. It is a fact generally recognized that in the first investigations on memory the amount of forgetting was unusually high, but modern investigators have found that the tenacity of the connections formed is more lasting than it was previously thought.

Figure 9 represents the forgetting curve for the four methods over a period of forty days. The parallelism of the lines illustrates the fact indicated previously of the consistency of the results. Let us note, however, that the curve for the method of context does not present the original drop as in the others. This is quite an exceptional curve in this respect. Let us notice also the almost horizontal line between the second and tenth days recalls.

Conclusions

1. The method of learning the vocabulary by associated pairs is superior to the method of learning it through a context or by a mixed method.

2. That superiority is very marked both for the immediate recall and for the delayed recall.

3. The method of associated pairs gives a great number of perfect scores, in fact two-thirds of the total as against one-third for the context method in a recall after fifty minutes.

4. When two consecutive recalls are made, the second, contrary to expectation is very little superior to the first. The only case where any noticeable superiority occurs is in the final recalls, after an interval of forty days.

5. There is a marked similarity in the results obtained after an interval of ten days, although no relearning has taken place.

---

[1] Brown, Warner: Effects of Interval on Recalls. *Journal Experimental Psychology*, 1924, Vol. VII, pp. 469–474.

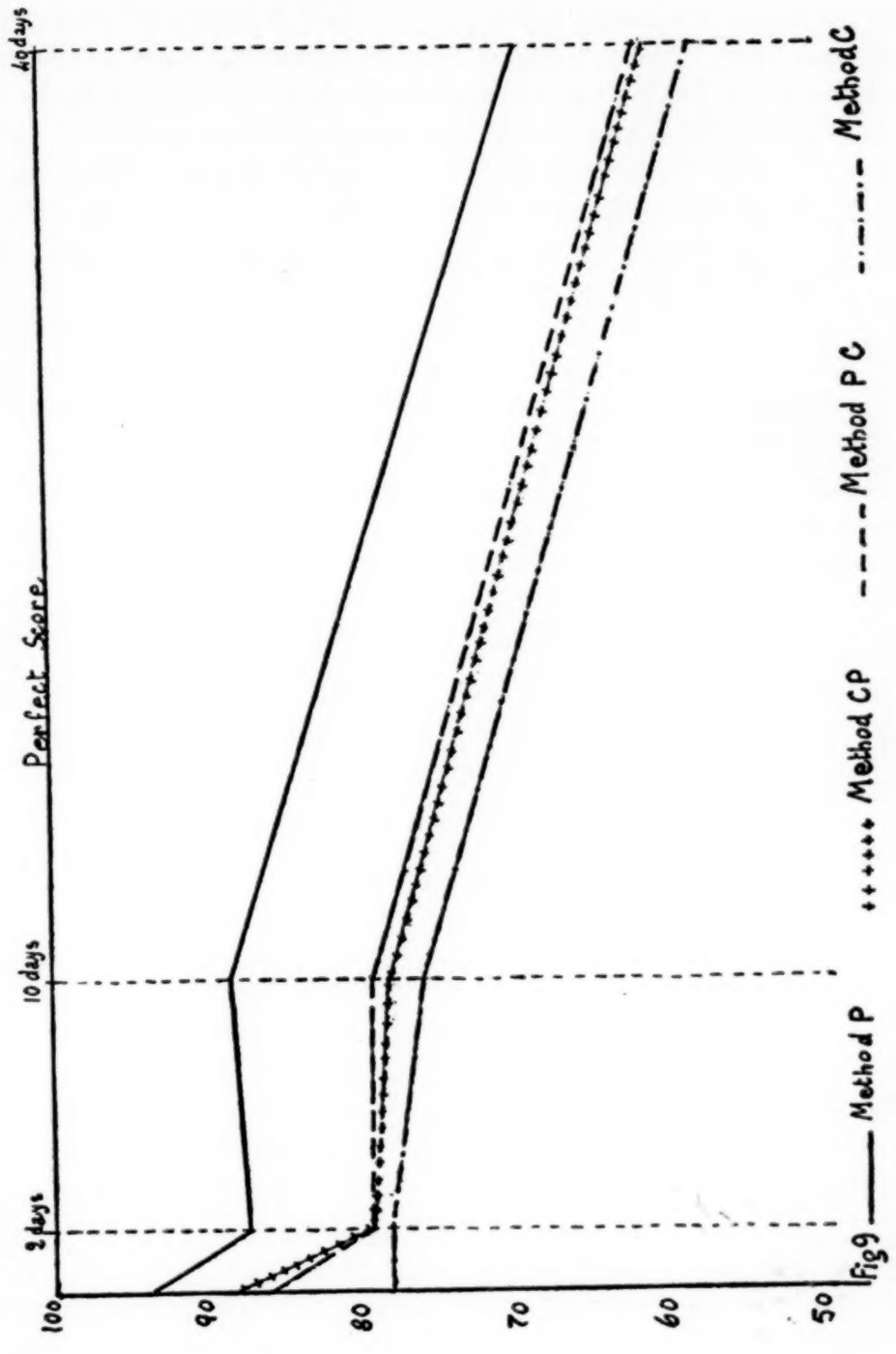

40 days
Perfect Score
10 days
2 days
100
90
80
70
60
50
Method C
Method P C
Method C P
Method P

Fig 9

6. An amount varying from ninety-three to sixty-nine per cent of the perfect score is retained after fifty minutes, and from seventy-nine to fifty-nine per cent of the perfect score is retained after forty days; this is in our estimation an unusual amount.

Correlations

We have calculated the correlations between the accuracy for each method (P recall) after an interval of fifty minutes. The results are as follows:

Correlation between methods P and PC	.50
Correlation between methods P and CP	.59
Correlation between methods P and C	.35
Correlation between methods PC and CP	.44
Correlation between methods PC and C	.50
Correlation between methods CP and C	.73

The correlations are all positive and fairly high except for Method P versus Method C. It is to be noted, however, that only seven students out of sixty actually did better by Method C than by Method P, according to the distribution diagram (steps of three units) and out of the seven, four made scores which differed by one step only (three units). The correlation between CP and C is the highest, .73.

The correlations between the immediate and delayed recalls are as follows:

For Method P	.49
For Method PC	.61
For Method CP	.50
For Method C	.56

The correlations are all positive and rather high. Compared with those obtained in another experiment they tend to show that these relatively higher correlations may be due to the fact that the amount retained in the final recall was so much higher in this experiment than in the other.

However, we may remark that the correlations between immediate and delayed recalls are as high as the correlations between the immediate recalls for the different methods of learning.

General Conclusions

It may be safely concluded that the method of learning vocabulary by associated pairs is by far the most superior of the four methods

tried in the experiment. This superiority maintains itself, through a period of forty days and for several recalls.

These results are by no means in accordance with the modern creed for the teaching of foreign vocabulary. There are schools which will not allow the use of any English at all in the teaching of a foreign language. In the reaction against the old methods of teaching foreign languages, we have gone so far as to discard everything, the good and the bad. However, a healthy reaction is beginning to be seen. Assertions made without the shadow of a proof or on the finding of "scientific" experiments which could not support scrutinity are beginning to be challenged. It is really comforting to see men like Palmer[1] taking the courageous stand of giving to each its proper due. Without denying the real advantages of the direct method, he says: "The exclusion of translation as a regular means of conveying the meaning of units is an uneconomical and unnatural principle."

Whatever may be said for the advantage of presenting the vocabulary in context (and there is much of course) still when it comes to memorizing the words as such, the method of associated pairs is the best. These results seem to justify the psychological law of specific learning. Learning is made of an infinity of small reactions which combined give a more complex reaction. The smaller the bond to remember, the better and quicker it will be remembered. Gates says: "One must learn precisely the reactions that will be needed for practical purposes."[2] It is evident that in using a word in a thousand possible combinations, the reactions obviously are limited in every case to the meaning of the word, and the context in which it was originally learned quickly fades out. Then why learn it at first?

In this experiment, when the students were warned in advance of the bond they were to remember, the presence of a text had very little influence; on the contrary, it proved to be a handicap in the formation of the bond between the English word and its French associate. That the students were able to use a word learned by the translation method in a context which they had not seen previously, answers the objection very often put forward that words learned that way are useless since the students are unable to use them when needed.

That the associations of a word with the context will help to strengthen the bond has also proven a fallacy.

---

[1] Palmer: "The Scientific Study and Teaching of Languages." P. 93.

[2] Gates, A. I.: "Psychology for Students of Education." P. 269.

This, of course, does not by any means imply that the study of the vocabulary in context ought to be avoided. It is perfectly evident that something very valuable is gained by such a process. For instance, there are French words that can not be satisfactorily translated by any English word. If a word has several meanings, a sentence to illustrate each meaning is also desirable. But the discussion here is not of the superiority of one method over the other in *toto*. It is specifically this: Given a certain number of French words to be studied to the extent of being able to give them either in a sentence or alone when the English word is given, the best method of study is by the method of associated pairs.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY FRANCES M. FOSTER

The Measurement of Nervous Habits in Normal Children, by Willard C. Olson. Univ. of Minn. Press, 1929. Pp. XII + 97.

The main contribution of the author is the development of a technique for observing a group of about thirty children and recording for each child, particular habits (*e.g.*, oral habits; sucking thumb, sucking fingers, biting nails, protruding tongue) in terms of number of five-minute periods in a series of twenty, during which at least one such moment was observed for that child. Oral habits were most commonly observed, with nasal, hirsutal, manual, irritational habits next in order, and aural, ocular, genital or facial movements less commonly observed. Reliability of the five best measures ranged from .45 for nasal to .73 for oral habits, with a mean intercorrelation of .34, which, corrected for attenuation would become .56. The correlation of oral habits with the remainder was .77. Records made on 467 children in Grades I–VIII indicate slightly more such behavior in girls, no relation to age, sibling correlations .19; twins .65, slightly more such behavior with known nervous defect in family line, some relation (r = .22) between such behavior and that of children seated for several months near and in front of the pupil, possibly more such nervousness with extremes of breast feeding, both more and less, slight tendency to more movements in afternoon than in morning, and slight tendency to more such movements among overweight or underweight children. The zero correlations with Woodworth-Mathews, Pressey X-O, and Kent-Rosanoff-Woodrow-Lowell Associations Tests are interpreted as due to the intellectual element in all of those tests. Curvilinear relationships were common. Oral habit scores and intelligence showed no relationship (−.01 for each age, 9–12).

Interesting techniques were devised, also for measuring eye-tremor, speed of tapping, and over-bite. The latter as a measure of thumb sucking seems related to present oral habits. An experiment upon two rats, with two others as controls, suggested the possibility

of persistence of movements of legs to head and ears, begun by collodium on the ear, for at least a month after the application.

GOODWIN WATSON.

Teachers College, Columbia University.

Handbook of Teaching Skills by W. H. Lancelot. New York: John Wiley and Sons, Inc., 1929. Pp. XII + 198. $2.00.

The title of this book is misleading. One expects to find a manual on how to teach, with the emphasis on concrete procedures, practical devices, and hints on classroom techniques. The emphasis, however, is on psychological principles rather than upon actual teaching skills. The author in the preface re-defines the term teaching skills in the following words: "The suggested procedures whereby learning by thinking may be substituted for the traditional process of learning by repetition and recall are referred to in this book as teaching skills."

The preface also states that the book "Represents in a sense a protest against the view too commonly held that intelligent human beings learn as the lower animals do; *i.e.*, by sheer repetition until stimulus and response, as we say, are joined together." In opposition to this view, the author advances "the three-fold proposition that the true learning process is essentially a thinking process."

The author deals with fifteen different teaching skills, as he calls them, which are treated at intervals along through a total of twenty one chapters. The specific skills treated are as follows:

1. Techniques for making external connections.
2. Replacing memorization with interesting thinking.
3. Involving the natural impulses.
4. Making internal connections.
5. Creating a feeling of need.
6. Creating and maintaining suspense.
7. The development of permanent interest.
8. Developing thinking ability in students.
9. Finding and using problems.
10. Putting the problem before the class.
11. Planning for the class discussion.
12. Leading the class discussion.
13. Formulating and putting questions.
14. Selecting the fact material to be taught.
15. Measuring the results of teaching.

Each of these is usually begun at the first of a chapter and is followed by about a page or two of rather practical suggestions for the classroom teacher. The remainder of the chapter then deals with various general considerations of a theoretical sort not in very close harmony with the practical nature of the page or two of introduction.

To the reviewer the book seems to be something of a cross between a text in psychology and a text in methods without being either in a complete and well rounded manner. Many practical teaching problems are not discussed. Many psychological principles are hardly mentioned. Some of the psychological principles that are mentioned are left without any very definite applications to teaching procedures.

The reviewer feels that the book is not constructed on a plan which will give it very wide applicability in the training of teachers but that it may be of considerable service in certain special situations. Its value would have to be judged in part in relation to the other courses which the student took in preparation for the teaching profession.

C. C. Crawford.

University of Southern California.

Ethics, by Frank Chapman Sharp. New York: The Century Co., 1928. Pp. 566.

The past few years have brought us an avalanche of books on ethics from the pens of professional and lay moralists. If a prize were to be offered for the most refreshing and significant material on the subject there is hardly any doubt that the reward would go to the lay writers. The output of the professional teachers of ethics is still musty and lifeless. Compared with writings such as those of Lippmann, Russell, Krutch, Ellis, Edman, and the late Sherman, the books of the academicians are dry and irrelevant. The one outstanding exception to this classification is Professor Fite's Moral Philosophy, a work of exceptional insight, vitality, and literary quality. Such a work is as rare as it is welcome.

Professor Sharp's book strikes a happy medium between the professional and lay writer. In a fresh and clear style he summarizes for the beginning student the problems and issues of ethics, and illustrates them with a happy choice of anecdotal material. His distinction between the idea of the *right* and that of the *good* is an

excellent point, striking at the very roots of absolutistic and individualistic morality.

The least that must be said in praise of Professor Sharp is that he has humanized and vitalized ethics for the college student, which is saying a good deal. His book will be welcomed by teachers of the subject, while students will read it with much pleasure and profit.

MAX SCHOEN.

Carnegie Institute of Technology.

The Will to be Free, by Howard B. Knox. New York: E. P. Dutton and Co., 1929. Pp. XVI + 237.

"The Will to be Free," written by "Captain Knox during his tenure of a research fellowship in philosophy at Manchester College, Oxford," bears all the ear-marks of work performed by an amateur under rigid thesis requirements. It is pedantic, heavily documented, and over-organized. Containing only 237 pages, it is still divided into seven parts, forty-two chapters, and countless sections. Such organization, instead of furthering clarity, irritates and confuses the reader.

With true clerical zeal Knox maintains that the deterministic principle can flourish only "in dark and noisome atmosphere." This phrasing is in itself indicatory of the rather excited attitude of the writer. The book, as a whole, lacks poise. It is difficult to extract from the mass of philosophic verbiage and Latin, French, Greek, and German phrases the main points of the author's arguments. However, two of them at least seem to be: First, that determinism is self-contradictory; and second, that it claims its arguments to be indisputable thereby immediately nullifying those arguments. Or, to quote Knox exactly: "A truth which it is impossible to doubt is presumably a lie." It is with such philosophical twisters that Knox attempts to prove the will free. The writer has added little, if anything, that is valuable to the reams of argument which have already appeared in connection with the determinism-indeterminism controversy.

HERBERT A. CARROLL.

Teachers College, Columbia University.